VOLUME 21, 1959/60, No. 7 pp. 185-220 Published 7th June 1960

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

SEGREGATION AND DISTRIBUTION OF IMPURITIES IN THE PREPARATION OF GERMANIUM AND SILICON *)

by J. GOORISSEN.

548.4:669.783:669.782

In connection with solid-state research and the production of circuit devices based on semiconductors, much work has been done in recent years on developing and perfecting methods of refining the substances involved and introducing controlled amounts of impurities. The article below describes two such methods (variants of the zone-melting technique) for preparing "doped" single crystals of germanium and silicon. These methods, evolved in the Philips Eindhoven laboratories, yield a product in which the concentration of the impurity is uniformly distributed.

In the fabrication of transistors and associated semiconductor devices it is necessary to be able to prepare the materials germanium and silicon with a high degree of purity, and also to "dope" them in a controlled way with a particular impurity. Further, the product is required in the form of a single crystal and must meet certain demands as regards physical perfection.

In all the various processes the product is obtained by the solidification of molten material. This involves the phenomenon, which has favourable as well as unfavourable aspects, that the impurity concentration in the solid at the freezing point differs from that in the liquid. If the concentrations are small, their distribution coefficient k_0 (i.e. the ratio of the concentration C_S in the solid to the concentration C_L in the liquid) has a constant value that usually differs very considerably from unity. For example, in the case of a solution of indium in germanium k_0 is only 1.4×10^{-3} (*fig. 1*).

In the following we shall first consider the consequences of this phenomenon when a dilute solution freezes progressively from a given point, i.e. when the solid-liquid interface moves through the charge. We shall see that, where $k_0 < 1$, a large fraction of the impurity gathers in the last part to solidify, so that the concentration in the remainder is reduced. This segregation process can therefore be used for purification. In the solidified material, however, the impurity concentration is not uniform, and from the point of view of doping, where the aim is to obtain a homogeneous final product, this is a troublesome effect. In the latter half of this article we shall discuss two doping methods developed in this laboratory, whereby special measures are taken to ensure

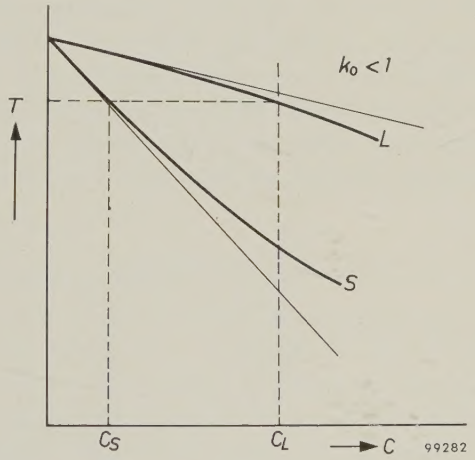


Fig. 1. In a two-component system the concentrations at a given temperature in the liquid and solid phases are not equal. When a liquid of concentration $C = C_L$ is cooled, crystallization occurs as soon as the temperature drops below that corresponding to C_L on the liquidus curve L. The concentration C_S in the resultant solid is the abscissa value of the solidus curve S at that temperature. For small concentrations S and L may be replaced by straight lines, and the ratio C_S/C_L is equal to the distribution coefficient k_0. If the two lines drop with increasing concentration, then $k_0 < 1$; if they both rise, then $k_0 > 1$.

*) This article is based on a lecture given 18 April 1959 at Moll (Belgium) before the Int. symp. on pure metals and semiconductors. The text of the lecture will shortly be published in the Vlaams Chemisch Weekblad.

that the impurity concentration is constant throughout a large part of the volume of the product.

Segregation and distribution of impurities in the normal solidification process; crystal pulling

In considering exactly what happens at a moving solidification boundary we shall suppose that the material is contained in an oblong boat in a furnace and is wholly molten. The boat is now slowly withdrawn from the furnace, so that the melt solidifies at one end. In relation to the furnace the solid-liquid interface remains at the same position, but in relation to the boat it moves from one end to the other. We assume that the solid-liquid interface is perpendicular to the length of the boat, which allows us to treat the problem as one-dimensional; the length of the boat represents the x axis. The origin is the place where the solidification begins. The x coordinate of the solid-liquid interface we shall call s (*fig. 2*). As long as the whole charge is

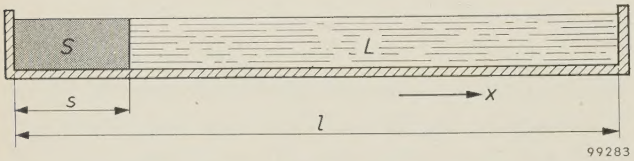


Fig. 2. Progressive solidification in a boat of length l. The x axis is the long axis of the boat. The solid-liquid interface is at the location $x = s$ and moves to the right.

molten, the concentration of the impurity is everywhere the same; let this value be C_i. In the material first to solidify, the concentration C_S is equal to $k_0 C_i$. For $k_0 \ll 1$, this solid part is thus appreciably purer than the liquid. The amount of impurity expelled from the volume which has frozen is now contained in the liquid. The further variation of C_S with increasing s depends on the manner in which this displaced amount of impurity is distributed over the bulk of the liquid. In principle, two limiting cases are conceivable: *a*) due to vigorous stirring, the amount of impurity taken up is immediately distributed uniformly throughout the bulk of the liquid, and *b*) there is no convection or stirring, so that the impurity is transported in the liquid solely by diffusion.

In the first case, for s fairly small relative to the length l of the boat, the concentration C_S in the solid increases only slowly with x; it has only double the $x = 0$ value at points $x \geqq l/2$.

In the other case, a concentration distribution arises that depends on the diffusion constant D and the rate f at which the solid-liquid interface moves, such that C_L *at* the interface is equal to C_i/k_0, and C_S is equal to C_i. This is illustrated in *fig. 3a*. Further from the interface, C_L still (in general) has the value C_i; the distance between the interface and the place where $C_L(x)$ differs from C_i by less than, say, 1% is denoted d in fig. 3*a*. When $D/f \ll l$, this "length" d of the diffusion region is for a long time smaller than the length $l - s$ of the liquid. Assuming, then, that for $x > s + d$ the liquid has exactly the concentration C_i, and since in the steady state as much impurity must enter and leave the diffusion region via the solid-liquid interface as enters and leaves it from the other side, evidently $C_S = C_i$. Except at the ends of the bar, C_S is thus equal to C_i in this case.

In reality, convection is not completely absent; there is, however, a boundary-layer of liquid at the advancing interface in which convection in the x-direction is negligible and in which the impurity can therefore be transported only by diffusion. If the circumstances — we shall turn to this point presently — are such that the thickness δ of this layer is less than the above-mentioned diffusion length d, the diffusion pattern will differ from that

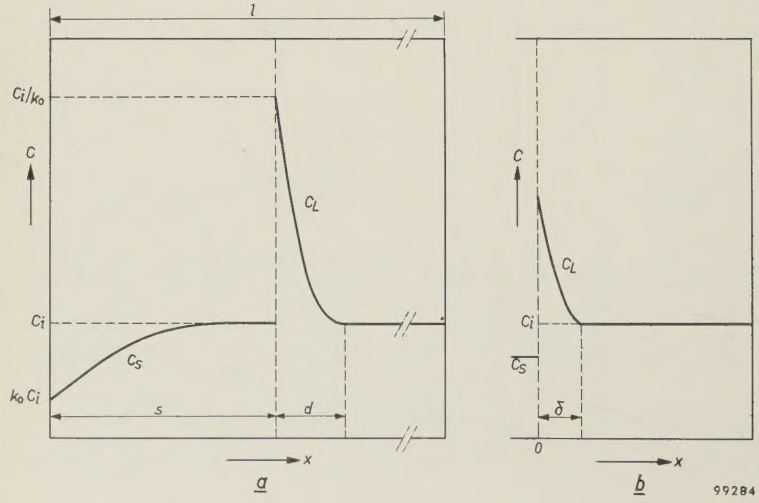


Fig. 3. *a*) If the transport of the impurity not taken up in the solidified material is due solely to diffusion, the diffusion process at the solid-liquid interface reaches a steady state as soon as C_L at the interface has risen to the value C_i/k_0.
b) The actual situation at the solid-liquid interface can be described by assuming the presence at this interface of a layer of thickness δ in which the impurity is transported in the x direction solely as a result of diffusion, whilst in the bulk of the melt convection occurs. When $\delta < d$ (fig. 3*a*), C_L at the solid-liquid interface does not attain the value C_i/k_0, and k ($= C_S/C_i$) is therefore smaller than 1. (In connection with the derivation of equation (1), the origin of the coordinate system is situated here in the interface.)

in the case without convection. The result is that C_S at the interface cannot reach the value C_i/k_0. The segregation coefficient k, which is equal to C_S divided by C_L measured at some distance from the solid-liquid interface, therefore has a value in this case between k_0 and unity (fig. 3*b*). It can be shown that this value is given by the formula [1]):

$$k = \frac{k_0}{k_0 + (1-k_0)\, e^{-f\delta/D}}. \quad . \; . \; . \quad (1)$$

Taking a coordinate system whose origin is in the solid-liquid interface (and which thus moves with the advancing interface), we can write (cf. fig. 3*b*):

$$D\frac{d^2C}{dx^2} + f\frac{dC}{dx} = 0 \quad . \; . \; . \; . \; . \; . \quad (2)$$

(omitting the subscript L), the boundary conditions being:

$$f\{C(0) - C_S\} + D\left(\frac{dC}{dx}\right)_{x=0} = 0, \quad . \; . \; . \quad (3)$$

$$C(x) = C_i \text{ for } x \geqq \delta. \quad . \; . \; . \; . \; . \; . \quad (4)$$

The solution is:

$$C(x) = C_i \frac{k_0 + (1-k_0)\, e^{-fx/D}}{k_0 + (1-k_0)\, e^{-f\delta/D}}. \quad . \; . \; . \; . \quad (5)$$

Putting $x = 0$, we find from this expression the concentration $C(0)$ at the interface, and multiplying this by k_0, we find the concentration C_S in the solid. Dividing the expression thus obtained by C_i, we finally arrive at the equation (1) for k.

If we substitute in (1) the value of k under given conditions, we can find the magnitude of δ from this expression. For antimony in germanium, δ is found to be of the order of 10^{-2} cm at an f value of 1.8 mm/min. (The case in which mass transport is due solely to diffusion can also be expressed by (2), (3) and (4): δ is then put equal to infinity in the second boundary condition (4). The solution — represented graphically by the curve in fig. 3*a* — is found from (5) by again putting $\delta = \infty$; the denominator of (5) then equals k_0. From this solution we can deduce that the distance d from the solid-liquid interface, where $C(x)$ differs from C_i by not more than 1%, is approximately 1.7 mm in the example of Sb in Ge.)

A plot of C_S versus x is given in *fig. 4* for the actual situation (curve *3*) as well as for the two hypothetical limiting cases. Curve *1* applies to the case where there is complete absence of diffusion but very vigorous stirring. Curve *2* applies to the case of no stirring. Except near the head of the sample, where the diffusion process has not yet reached the steady state, curve *3* shows the same general form as curve *1*. It can be seen that, although a substantial fraction of the impurity is displaced towards the last part to solidify, everywhere in the sample the value of C_S varies somewhat with x. This means that although the method of solidification described above is suitable for purposes of purification — and, as such, is comparable with the re-crystallization methods used in chemistry — it is less suitable as a method of doping.

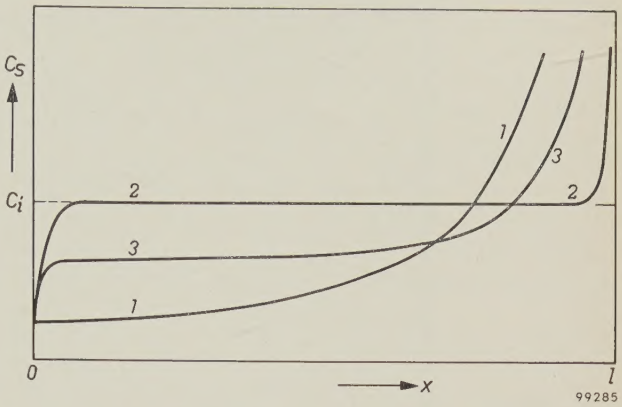


Fig. 4. The variation of C_S with x in ingots produced by the progressive solidification from one end of a melt having a concentration C_i. Curve *1*: the melt is vigorously stirred. A large portion of the ingot is highly purified, but C_S does not have a constant value. Curve *2*: no convection in the melt, the impurity being transported solely by diffusion. Over a considerable length of the ingot $C_S = C_i$, i.e. k is constant over this region and equal to unity. In reality these extremes do not occur, and C_S follows curve *3* ($k_0 < k < 1$). Except where x is very small, curve *3* has the same character as curve *1*.

Except at very small values of x, the variation of C_S with x is given for all possible situations by:

$$C_S(x) = kC_i(1 - x/l)^{k-1}. \quad . \; . \; . \quad (6)$$

In the first limiting case mentioned above (fig. 4, curve *1*) we had $k = k_0$, and in the other case (curve *2*) we had $k = 1$. In reality, of course, the value of k lies somewhere between k_0 and 1.

Finally, it should be noted that $C_S(x)$ can be indirectly determined by measuring the local resistivity of the material. The relation between the two quantities is known: the conductivities of impure germanium and silicon are to the first approximation proportional to C_S.

The pulling of single crystals

A widely used method of making single crystals is the Czochralski technique [2]). A seed crystal attached to a rotating vertical shaft is lowered until it just touches the surface of a quantity of molten material. The temperature of the melt at that position must be just equal to the freezing point in order to prevent the seed crystal from melting away. The shaft with the seed crystal is now slowly raised, and at the same time the heat supplied to the melt is slightly reduced. As soon as a steady state is reached, a rod-shaped single crystal

[1]) J. A. Burton, R. C. Prim and W. P. Slichter, J. chem. Phys. **21**, 1987, 1953.

[2]) J. Czochralski, Z. phys. Chem. **92**, 219, 1917.

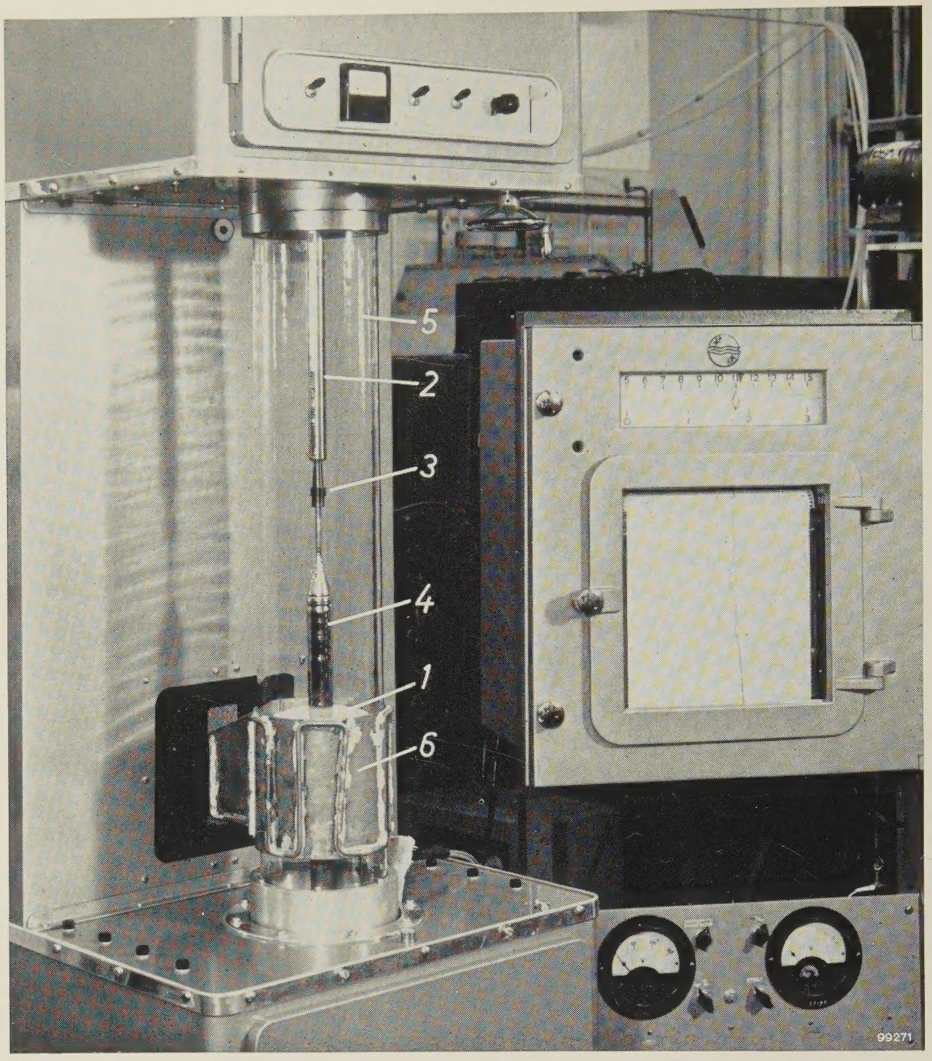


Fig. 5. Apparatus for pulling a single crystal from the melt (Czochralski method). *1* crucible containing molten germanium. *2* rotating shaft. *3* seed-crystal holder. *4* pulled single crystal. Components *1* to *4* are contained inside a glass envelope *5* through which an inert gas flows. The high-frequency heating is provided by a coil *6* in the form of a water-cooled plate encircling the glass envelope at the position of the crucible. The automatic potentiometer (right) holds the crucible at the desired temperature by automatic control of the heating supplied by the high-frequency generator; it also records the actual temperature.

of constant diameter grows on the seed crystal (*fig. 5* and *6*).

The situation at the solid-liquid interface in this case is similar to that in the solidification process described above, in which the liquid was stirred except in a layer of thickness δ. The magnitude of δ is here determined by the viscosity of the liquid and the speed of rotation of the shaft. The value assumed by the segregation coefficient k is therefore determined here by these quantities as well as by the speed at which the crystal is pulled and by the diffusion constant.

It may be concluded from the above considerations that, here too, C_S is nowhere in the single-crystal rod independent of position. Since the pulled crystal is the final product of the preparation process, this is a very undesirable state of affairs.

The equations given earlier [1]) for the solidification process were in fact first derived for the method of crystal-pulling just discussed. It should be noted here that the solution (5) is in fact an approximation, since it is an over-simplification to assume that there exists in the liquid a layer of thickness δ in which diffusion is alone responsible for impurity transport and which is sharply divided from the remainder. If the problem be posed exactly (that is to say for the pulling method with rotating crystal), it can only be solved numerically. For the point $x = 0$ the solution can be written in the form of an integral equation, but it is impossible to derive from this a formula for k.

The validity of the approximate treatment given above is corroborated by an experiment based on the following considerations. Since a single crystal is required, the temperature $T(x)$ in the liquid must everywhere be higher than the equilibrium temperature T_E which can be read for a given concentration from the phase diagram. If this condition is not satisfied, multiple nucleation will occur. Assuming that there is no supercooling — at the interface the temperature is necessarily equal to T_E — the requirement $T(x) > T_E$ in the immediate proximity of the interface ($x = 0$) can be replaced by:

$$\left(\frac{dT}{dx}\right)_{x=0} > \left(\frac{dT_E}{dx}\right)_{x=0}. \quad \ldots\ldots \quad (7)$$

The right-hand side of (7) can be written:

$$\left(\frac{dT_E}{dx}\right)_{x=0} = \left(\frac{dT_E}{dC_L}\right)_{C_L=C(0)} \times \left(\frac{dC_L}{dx}\right)_{x=0}. \quad \ldots \quad (8)$$

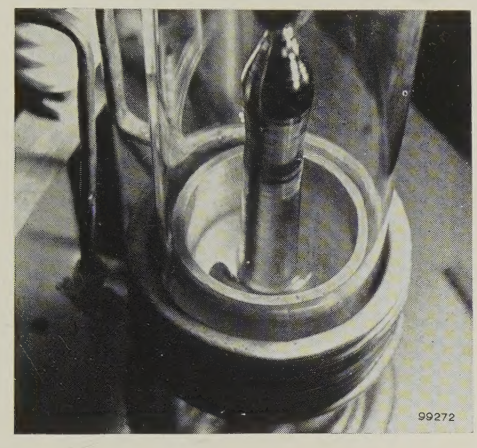


Fig. 6. Pulling a single crystal from the melt. The operation is almost complete. The photograph shows clearly the shape of the crystal obtained.

From (3) it follows that

$$\left(\frac{dC_L}{dx}\right)_{x=0} = -\frac{f\{C_L(0) - C_S(0)\}}{D} = -\frac{f}{D} C_L(0)(1-k_0), \quad (9)$$

so that

$$\left(\frac{dT_E}{dx}\right)_{x=0} = -\left(\frac{dT_E}{dC_L}\right)_{C_L=C(0)} \times \frac{f}{D} C_L(0)(1-k_0). \quad (10)$$

Substituting in this expression the values of dT_E/dC_L, D and k_0 for a given impurity and the relevant value of f, and putting dT_E/dx equal to the value found for dT/dx, we can calculate from (10) the maximum value that $C_L(0)$ may have under the given circumstances without the occurrence of nucleation.

In the case of germanium containing antimony as impurity, we have:

$$\frac{dT_E}{dC_L} = 400 \text{ °C/mol},$$

$$k_0 = 0.003,$$

$$D = 5 \times 10^{-5} \text{ cm}^2/\text{sec}.$$

Suppose $f = 1.8$ mm per minute or 3×10^{-3} cm/sec. Assuming that dT/dx in the solid is about 150 °C/cm at the solid-liquid interface, we can draw up an equation between the rates of heat transport to and from the interface ($x = 0$); we thus calculate that dT/dx in the liquid is approximately 28 °C/cm (taking the latent heat of fusion as 98 cal/gram and the thermal conductivity coefficient of solid and liquid Ge as 0.034 and 0.12 cal/sec.cm.°C respectively).

For $C_L(0)$ this gives a value of 1.4×10^{-3} mole/cm^3, and hence for $C_S(0)$ a value of 4.2×10^{-6} mole/cm^3, that is to say 2.6×10^{18} atoms of antimony per cm^3. Calculation shows that the resistivity of the germanium with this value of C_S should be 0.004 Ωcm. Measured on a piece of germanium doped with Sb, the resistivity at the boundary between the single-crystal and polycrystalline region, i.e. at the position where T is still just above the value T_E, is found to be 0.003 Ωcm. Since the values taken for dT/dx and for the thermal conductivity are not known exactly, and since some supercooling occurs before nucleation begins [3]), the agreement is highly satisfactory.

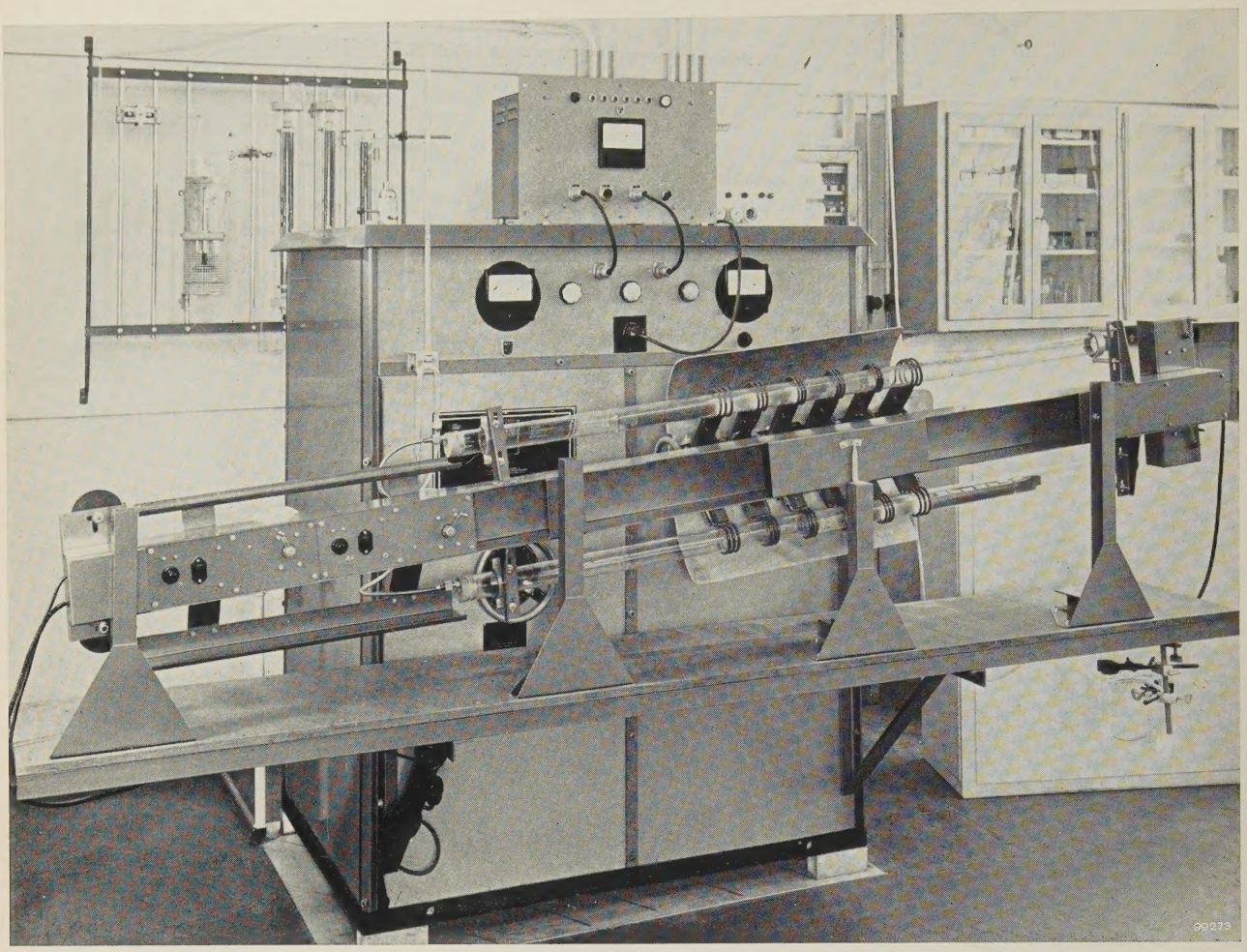
Modern doping and purification methods

The modern methods of doping germanium and silicon are all variants of the crystal-pulling method or of zone melting; purification is always effected by a zone-melting process. In the latter the ingot is locally heated so as to melt only a short zone of its length. The heating is usually inductive with the aid of a high-frequency generator, the molten zone being made to traverse the ingot slowly from one end to the other.

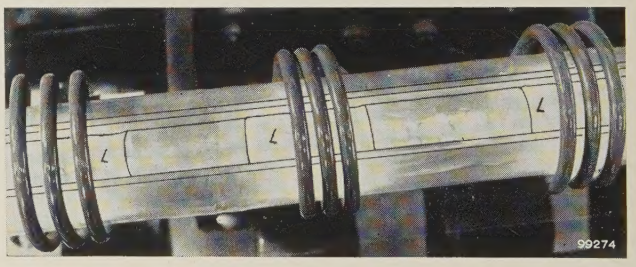
If we compare this process with that of causing one end of an entirely molten ingot to solidify, we notice that the amount of impurity segregated from the solidified part — we again assume $k_0 \ll 1$ — is taken up in a smaller volume of liquid, so that the increase in C_L is more pronounced and the purification less effective. Whereas in normal solidification C_L does not reach twice its initial value before half the ingot is solidified, in this case it reaches this value after only slightly more than two zone lengths.

The fact that zone melting is nevertheless preferable as a method of purification is due to the following circumstances. In the first place the process can be repeated indefinitely without it being necessary to cut off the part last to solidify, in which C_S is, of course, very large. Apart from the time saved, this is an advantage inasmuch as there is no risk of contaminating the ingot by any mechanical or manual handling. Further, a second and third zone can start to traverse the ingot before the first has reached the end. *Fig. 7* shows photographs of a zone-refining equipment for germanium in which the ingot passes through no fewer than six heaters, so that in a single run six molten zones pass through the ingot. This obviously gives an

[3]) B. Chalmers, Proc. Int. Conf. on Crystal Growth, Cooperstown (N.Y.) 1958, p. 297, Wiley, New York.



a



b

Fig. 7. *a*) Apparatus for purifying germanium by multiple zone melting. The graphite boats containing the germanium are enclosed in tubes through which dry purified hydrogen flows. During the zone-melting process the boats move slowly from right to left; in each run they pass six high-frequency heating coils situated at short distances apart. They are pulled forward by means of a thin metal wire. The tubes are on a slight slope to prevent the solidified charge from assuming a wedge shape due to the differing specific volumes of liquid and solid. At the end of a run the boats can be rapidly returned by mechanical means to their starting point in order to start a second run, and so on. The equipment on the wall rack is used for purifying the hydrogen.
b) Detail of the apparatus. The molten zones (L), at the position of the heating coils, are 4 to 5 cm long.

enormous saving of time. Nor is it necessary to handle the charge if it is to be passed through for a further run.

Since molten silicon is chemically highly reactive, melting it in a crucible gives rise to problems. Nevertheless it can be processed successfully by a special method of zone melting known as the floating-zone technique. In this method the charge is clamped at its ends in a vertical position. Provided it is not too long, the zone is held together by its own surface tension (*fig. 8*). Here, too, the growing crystal is made to rotate. A single-crystal product is obtained provided the end from which the molten zone begins is a single-crystal seed in contact with the charge proper. The floating-zone technique is the best method of making single-crystal silicon.

It may be inferred from what has been said that, in the forms discussed above, zone melting for the purposes of doping is open to the same objections as the methods outlined in the introduction. The concentration C_S is not constant over any appreciable length of the product.

An improvement is possible, however, by modifying the method. Where k is very small (<0.01)

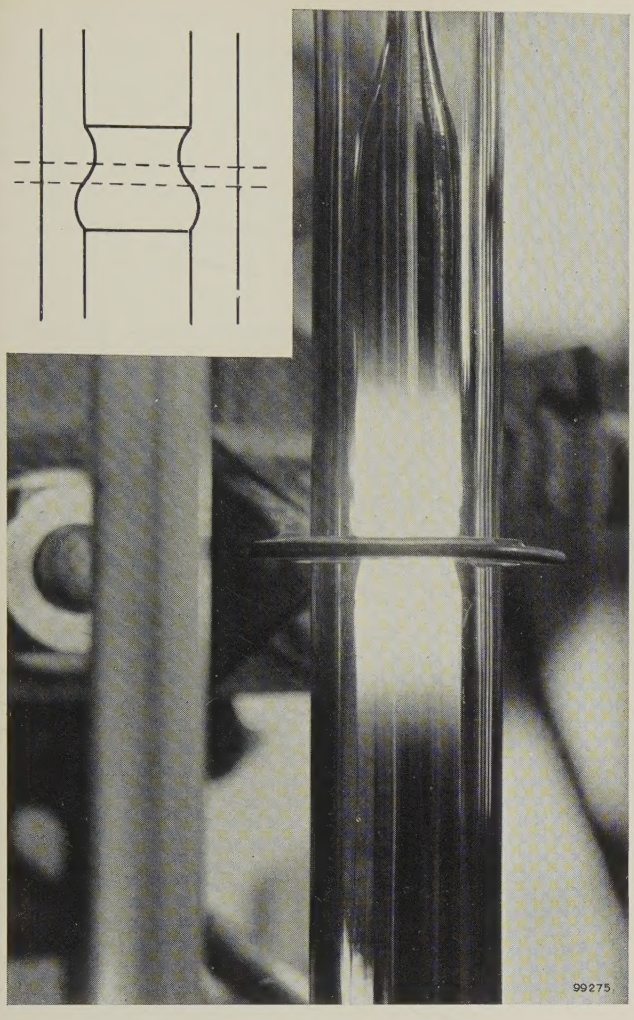


Fig. 8. Purification of silicon by the floating-zone technique. The charge is here set up vertically, clamped at the ends, in the middle of a relatively wide tube through which an inert gas flows. The (single) zone is held together by the surface tension of the molten material (see inset).

good results can be obtained, for example, by arranging for the first zone of the ingot to consist of impure material (concentration C_i) and the remainder of pure material. Except of course in the last zone, one then obtains approximately the concentration $k_0 C_i$ over the entire ingot. If k_0 is not very small, improvement is possible by causing a zone to pass to and fro many times, or by making use of the fact that C_S also depends on the length of the zone; in the latter case one aims at keeping C_S constant by changing the zone length during the course of the process. Use can also be made of the fact that k depends on the speed at which the zone moves through the ingot [4]). The first process is, of course, time-consuming, and all three are subject to practical difficulties.

The best doping methods are essentially those in which a highly-purified material is used and just as much of the required impurity is continuously supplied from outside to the molten zone as is withdrawn from the trailing solid-liquid interface. A product is then obtained in which C_S is entirely uniform throughout, irrespective of the value of k_0. We shall now discuss two processes developed in this laboratory which are based on this principle and have proved their usefulness. The first concerns the preparation of silicon doped with phosphorus [5]) — in this case $k_0 = 0.35$ and the method with the enriched first zone is therefore unsuitable — and the second the preparation of doped germanium [6]). The latter process can in principle be used for introducing virtually any impurity.

Phosphorus-doping of silicon by zone melting with continuous phosphorus feed

In the uniform vapour-phase doping of silicon with phosphorus the continuous phosphorus feed is effected by adding a trace of phosphine (PH_3) to the inert gas flowing around the ingot (the zone melting is done here of course by the floating-zone technique). The passage of the hot zone causes the PH_3 to decompose, and a certain fraction of the liberated P atoms is dissolved in the molten silicon. In the steady state the zone takes up as many P atoms per unit time as leave it via the trailing solid-liquid interface. The initial ingot of silicon must be well purified beforehand. The volume of the zone must remain constant during the process.

At the beginning of the process the ingot is surrounded only by a stream of inert gas. When a molten zone has formed, PH_3 is added at a constant rate to the stream of gas. As soon as the P concentration in the zone has attained the required value — we shall see presently how this is ascertained — the ingot (that is to say the zone, seen from the ingot) is set in motion.

Fig. 9 shows a plot of resistivity versus position for various single-crystal silicon rods produced in this way. It can be seen that over a very considerable length of each rod the resistivity deviates by no more than about 10% from the average value. The

[4]) The various methods of zone melting are discussed in W. G. Pfann, Zone melting, Wiley, New York 1958.

[5]) J. Goorissen and A. H. J. G. van Run, Gas-phase doping of silicon, Proc. Instn. Electr. Engrs. 1959, in press.

[6]) J. Goorissen and F. Karstensen, Das Ziehen von Germanium-Einkristallen aus dem „schwimmenden Tiegel", Z. Metallkunde **50**, 46-50, 1959 (No. 1), and J. Goorissen, F. Karstensen and B. Okkerse, Growing single crystals with constant resistivity by floating-crucible technique, published in Solid state physics in electronics and telecommunications, Proc. int. Conf. Brussels, June 2-7, 1958, edited by M. Désirant and J. L. Michiels, Vol. 1, pp. 23-27, Academic Press, London 1960.

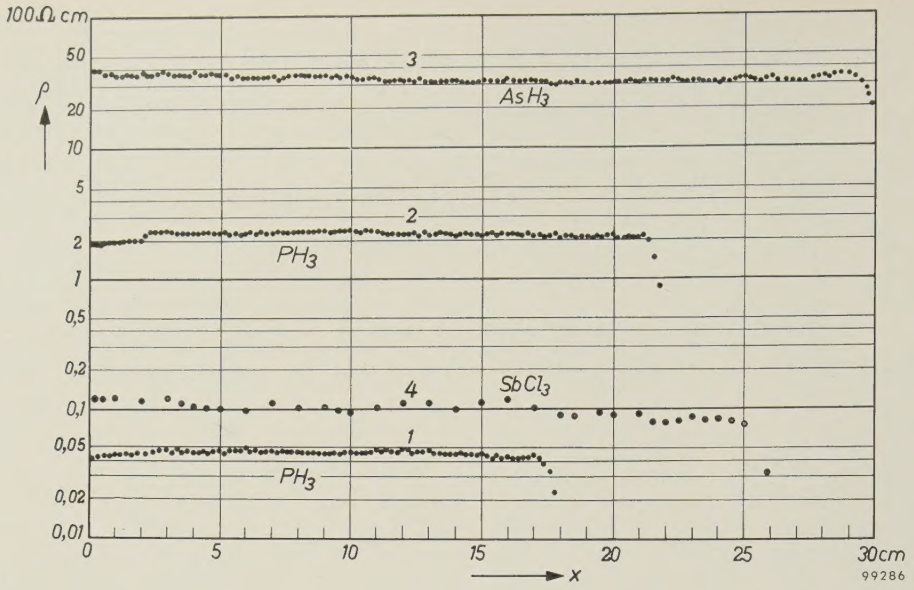


Fig. 9. Series *1* and *2*: resistivity ϱ as a function of position (x), measured on silicon rods doped with phosphorus; the impurity was added in the floating-zone treatment by means of a constant stream of PH_3. Except in the last zone, ϱ differs nowhere by more than 10% from the average value. Since for silicon containing phosphorus k_0 is 0.35, this degree of constancy of ϱ is very difficult to achieve by other methods. Series *3* and *4*: as above, but for silicon doped in the same way with arsenic (via AsH_3) and antimony (via $SbCl_3$), respectively.

same figure shows some provisional results of experiments using the same method for the doping of silicon with arsenic (via AsH_3) and with antimony (via $SbCl_3$).

An investigation into the principles governing the uptake of phosphorus in the molten zone has shown that, other conditions being equal (i.e. geometry, rate of zone travel and speed of rotation), the amount taken up per second in the growing rod depends solely on the amount of PH_3 flowing to the zone per second. If this amount is increased by, say, a factor of 2, the conductivity of the rod increases by the same factor; see *fig. 10*. It is immaterial in this connection whether the increase is brought about by doubling the PH_3 concentration with the gas flow remaining constant, or by increasing the gas flow with the concentration remaining constant. Finally, if we increase the speed of the zone, other conditions remaining equal, the phosphorus concentration in the rod decreases by the same factor. This means that the rate at which phosphorus is absorbed in the zone does not depend upon the rate at which the zone travels; the rate of absorption is also evidently independent of the phosphorus concentration in the zone at any instant. It may therefore be assumed that the rate of absorption during the initial period, in which the molten zone is not yet moving, is the same as when the zone is in motion. It is thus possible to calculate the time that must elapse, after introducing the PH_3 stream, before the zone can be started moving.

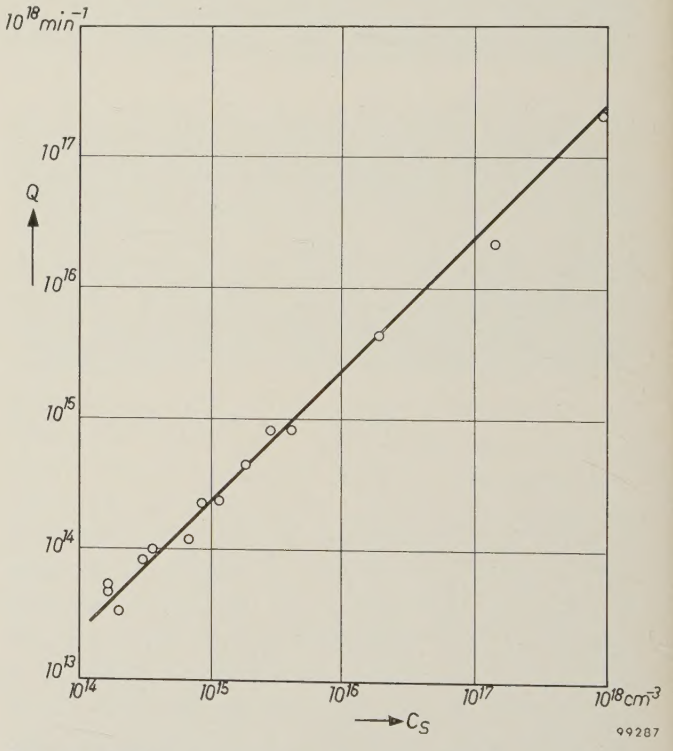


Fig. 10. Relation between the logarithm of the phosphorus concentration C_S in a silicon rod and that of the PH_3 stream Q to the zone. Over a very large region C_S and Q are related by a straight line of slope unity; the phosphorus concentration is therefore directly proportional to the quantity of PH_3 flowing per unit time. The units employed are P atoms per cm^3 for C_S, and PH_3 molecules per minute for Q.

A floating-crucible technique for the preparation of homogeneously doped germanium crystals

The floating-crucible technique [6]) for the preparation of doped single-crystal germanium with a homogeneous distribution of impurity is a variant of the Czochralski crystal-pulling method. Although it differs considerably from the floating-zone method just described, the basic idea is the same: in this case too a zone of constant volume is present to which as much impurity is added per second as leaves to enter the growing crystal from the solid-liquid interface. Again, the method is practicable whatever the value of k_0.

The method is illustrated schematically in *fig. 11*. The molten material *1* from which the single crystal *2* is pulled, is contained in a small graphite crucible *3* which floats on the melt *4* contained in a second crucible *5* mounted on legs. The contents of both crucibles are in communication via a capillary *6*. A stem underneath the floating crucible passes through a hole in the bottom of the large crucible and carries a weight *7*. The surface tension of germanium is sufficient to prevent the liquid from flowing away through the gap between the stem and the bottom of the large crucible.

The procedure is to fill the large crucible with pieces of pure germanium and to melt them, whilst the small crucible lies at the bottom. It remains there after the germanium is melted. The desired impurity is then added to the melt until its concentration is equal to the value C_S required in the product. With the spindle *8* the small crucible is now pushed up until it partly projects above the surface of the melt. The crucible then remains floating because the hydrostatic upward (buoyancy) force is now supplemented by an upward force due to the surface tension. Owing to the slope of the side walls of the small crucible, this capillary force varies with the depth of submergence. As a result vertical movements, due for example to external disturbances, remain small in amplitude.

Next, the rotating seed crystal is lowered on to the surface and the process of pulling a single crystal begins. This gives rise to a flow of liquid in the capillary. Thereupon an amount of the desired impurity is added to the germanium in the *small* crucible until the concentration there rises to C_S/k. Since the volume of the liquid in this crucible does not depend on the level of the liquid in the large crucible — we shall explain the reason presently — from that moment a stationary state is set up, provided the capillary is so narrow that the number of impurity particles leaving the floating crucible per second by diffusion is negligible compared with the number flowing in with the germanium. (It can be shown that this is the case if the rate of flow of the germanium in the capillary is high compared with the ratio of the diffusion constant to the length of the capillary.) Apart from a short region at the beginning and end, the single-crystal rod thus obtained possesses over its whole length a constant impurity concentration equal to the desired value C_S.

The constancy of the quantity of material in the small crucible is due to the fact that the stem carrying the weight *7*, and also the part of the crucible vertically above it, make no contribution to the hydrostatic upward force. This remains constant even though the liquid level drops in the large crucible. The submergence of the floating crucible does not therefore change as the operation proceeds; the volume of germanium therein remains constant.

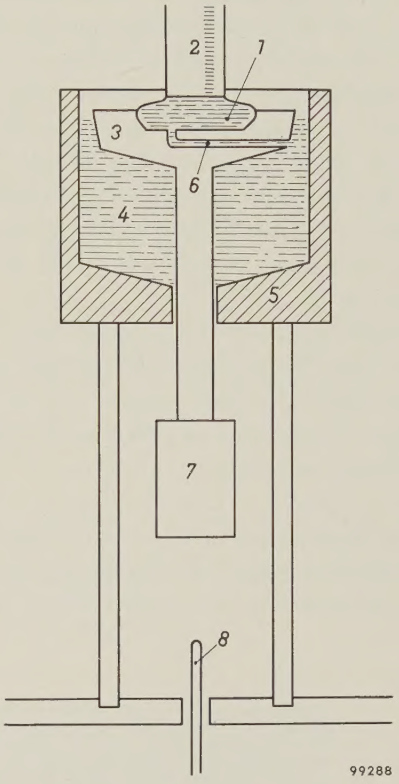


Fig. 11. Schematic cross-section of an arrangement for making single crystals of germanium with constant impurity concentration (floating-crucible technique). *1* molten germanium from which the crystal *2* is pulled and which is contained in the crucible *3*. This floats on a much larger quantity of germanium *4* with which the crucible *5* is filled. *6* connecting capillary. *7* weight. *8* spindle for pushing up the crucible *3* after the germanium in *5* is melted.

As might be inferred from the foregoing, the weight *7* is so chosen that the total weight of the small crucible, the stem and *7* is greater than the hydrostatic upward force. Consequently the small crucible not only remains at the bottom when the large one is being filled — this is necessary in order

to fill the capillary, though that could of course also be done without the weight 7 — but also the submergence when floating is sufficient. In the absence of 7 the small crucible would not sink deeply enough, and therefore it would lose all its liquid.

From the point of view of apparatus and procedure the method described here is closely related to the ordinary method of crystal pulling. Having regard, however, to the various volumes involved (those of the pulled crystal, the contents of the floating crucible and the contents of the large crucible) and considering the concentrations, it is evident that it can in fact be regarded as an unusual form of zone melting. The contents of the small crucible play the part here of an artificially enriched molten zone, and the large crucible can be thought of as the initial charge. The latter, however, already possesses the desired concentration (which is moreover independent of position), so that as the operation progresses the "zone" does not change in concentration.

Fig. 12 shows the result of resistivity measurements as a function of position for various germanium rods made by this method. Here, too, the resistivity in a large region differs by less than 10% from the average value.

Finally, a remark about the effectiveness of the two new methods discussed above. By this is meant the percentage of the total rod length in which the extreme resistivity values do not differ by more than, say, 20%.

Assuming that the concentration necessarily differs from the desired value only in the last zone, but is otherwise everywhere constant, the effectiveness of the two new processes is theoretically equal

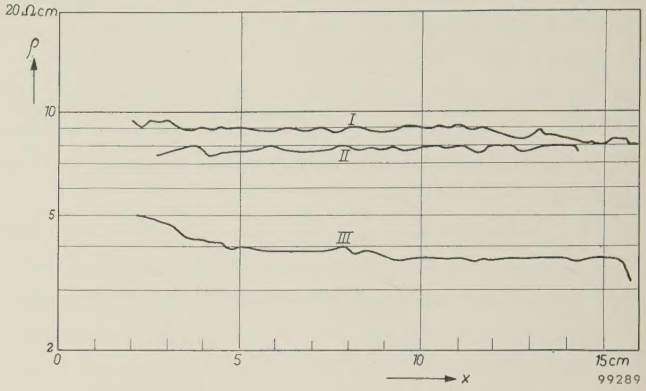


Fig. 12. Resistivity ϱ as a function of position (x), measured on various germanium crystals made by the floating-crucible technique. The impurities used were: curve I — indium ($k_0 = 0.0014$); curve II — antimony ($k_0 = 0.007$); curve III — phosphorus ($k_0 = 0.12$). Over most of the length of each crystal the resistivity ϱ is constant to within 10%.

to $100(1-V_z/V)$ %, where V_z is the volume of the zone (or of the contents of the floating crucible) and V is the total volume. In the floating-zone method, where silicon is continuously doped with phosphorus, this theoretical effectiveness, which is 90% for a rod ten zones long, can in fact be achieved. It is not quite realizable, however, with

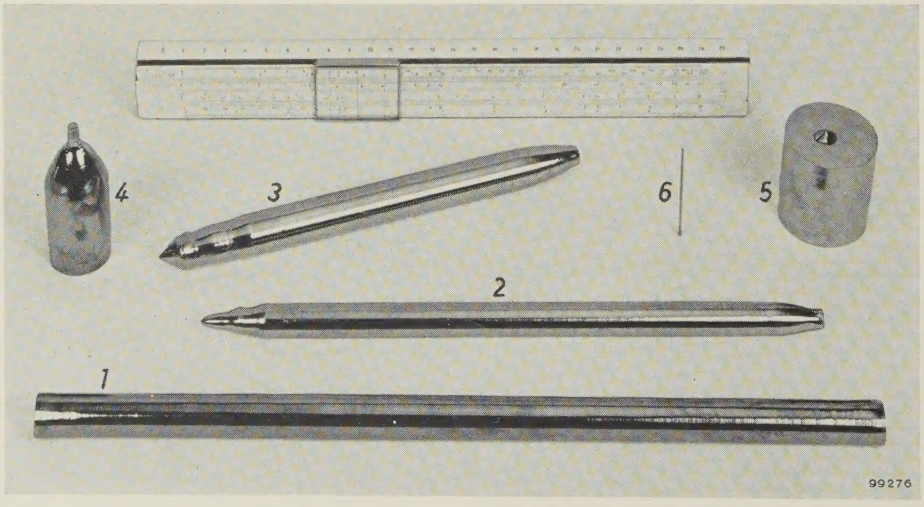


Fig. 13. Single crystals of various dimensions, made by the methods discussed in this article. *1* silicon crystal with constant phosphorus concentration (floating-zone melting with PH_3 doping), *2* and *3* germanium (floating-crucible method), *4* top end of a thick germanium crystal (floating crucible), *5* bottom end of a very thick germanium crystal (normal pulling method), *6* very thin germanium crystal (idem). From *1* and *2* it can be seen that the crystals do not have a circular cross-section. This is due to the fact that the rate of growth is not the same in all crystallographic directions.

the floating-crucible method. In the first place the volume of the liquid in the small crucible depends on the diameter of the rod (owing to the effect of the surface tension) and is therefore not constant at the beginning while the diameter of the crystal is still growing. In the second place, not all the germanium in the large crucible can be used. The effectiveness obtainable in practice with our apparatus is found to be about 75%; the theoretical efficiency is 85%.

Just as high an effectiveness can also be obtained with the earlier-mentioned zone-melting method using an enriched first zone, but only if k_0 is smaller than about 0.01; with increasing k_0 the efficiency drops sharply.

The effectiveness of the ordinary crystal-pulling method — we refer to germanium of course, silicon being handled without a crucible — is poor in the majority of cases. By varying the pulling rate, it can be raised to a theoretical value of 73% [7]) in the most favourable case (viz. with antimony, where k varies most strongly with the pulling rate). However, the complications involved are considerable.

Fig. 13 shows a photograph of a silicon rod and various germanium rods. All are single crystals; the first was obtained by the floating-zone technique; all but two of the germanium rods were produced by the floating-crucible technique.

[7]) Discussed at length in the first of the articles quoted under [6]).

Summary. At the solid-liquid interface of a substance containing a very small amount of an impurity, the concentration of this impurity in the solid phase generally differs from that in the liquid phase. The ratio k_0 between both concentrations is called the distribution coefficient. As the solidification progresses, the segregation constant k, i.e. the ratio between the concentration in the solidified part and that in the liquid (away from the solid-liquid interface), lies between k_0 and 1. If the liquid is vigorously stirred, then $k \approx k_0$; if there is no convection, then $k = 1$. The segregation process can be used for purification ($k < 1$), but it makes it difficult to obtain a homogeneous product. Nowadays, silicon and germanium are purified by the method of multiple-pass zone melting, whereby fairly short molten zones are passed from one end of an ingot to the other. In germanium, which can be melted in a crucible, the successive zones are passed through at short distance apart, and this represents a considerable saving of time. Silicon in the molten state is too reactive to be melted in a crucible. In this case one zone is passed through a vertical charge, clamped at either end, the zone being held intact by its own surface tension (floating-zone technique). The segregation phenomenon is troublesome if the material is to be doped. In general, a homogeneous product is not obtained (unless k_0 is very small and an enriched first zone is used). Two new methods have been evolved in the Philips laboratories. They are based on the idea that, irrespective of the value of k_0, a homogeneous product is obtained if as much impurity is added from outside to the zone per second as leaves it to enter the growing crystal via the solid-liquid interface. A homogeneous rod of single-crystal silicon, doped with phosphorus (k_0 is 0.35 for phosphorus in silicon), can be produced by adding a trace of PH_3 to the inert gas flowing around the charge. In the steady state the molten zone takes up as much phosphorus from the gas — PH_3 dissociates at high temperatures — as is expelled from it via the trailing solid-liquid interface. Homogeneously doped single-crystal germanium (impurity concentration C_S) can be made by the floating-crucible technique, which is a variant of the ordinary method of crystal pulling. The crystal is pulled from the contents of a small crucible, concentration C_S/k, which communicates via a capillary with the contents of a larger crucible of concentration C_S. As the operation progresses, the contents of the small crucible change neither in volume nor in concentration. The effectiveness of both methods as regards the homogeneity of the product is high and approaches close to the theoretical value, viz. $100(1-V_z/V)$ %, where V_z is the volume of the zone (or of the melt, in the floating crucible) and V is the total volume.

AN AUTOMATIC DEW-POINT HYGROMETER USING PELTIER COOLING

by P. GERTHSEN *), J. A. A. GILSING and M. van TOL.

537.322.15:536.423.45:662.613.54

Progress in the field of semiconductors has in recent years opened up more and more new applications. The article below describes an automatic dew-point hygrometer in which the necessary cooling is produced by the Peltier effect of a certain combination of semiconductors.

The most accurate method of measuring the humidity of air and other gases still remains the determination of the dew point. Instruments based on other commonly used methods are either dependent on material properties and must therefore be regularly calibrated (hair-hygrometers, diffusion method), or they are liable to get dirty, without the resultant errors becoming perceptible (electrolytic methods, wet and dry bulb). The other methods proposed, such as weighing before and after drying, measuring the speed of sound or the dielectric constant are scarcely suitable for industrial application.

A complication of the dew-point method is the cooling required in order to reach the dew point. For this reason the dew-point method has hitherto been used only in laboratories, and has not been adapted to industrial use.

As a result of intensive solid-state research in recent years, semiconductors are now available which make electrical cooling possible by very simple means, based on the use of the Peltier effect.

This effect (one of the group of thermoelectric effects) occurs when an electric current is passed through a circuit of two conductors of dissimilar material connected in series. Heat is generated at the one junction, whilst heat is absorbed at the other. The amounts of heat absorbed or developed per unit time are proportional to the current. The proportionality factor is the Peltier coefficient P of the particular combination of conducting materials. If the situation is reversed, i.e. if the circuit is opened and the junctions are brought to different temperatures, there appears across the terminals of the thermocouple so produced a potential difference that depends on the temperature of each of the junctions. The latter phenomenon, called the Seebeck effect or more generally "the" thermoelectric effect, is commonly used for temperature measurement [1]).

The apparatus for determining the dew point is represented schematically in *fig. 1*. Blocks *1* and *2*, made of different semiconducting materials, are joined at positions *a* by a silver plate *3*, about 5 mm thick and wide and about 15 mm long. At positions *b* they are connected to fairly large and thick copper plates which act as "cooling fins" and also conduct

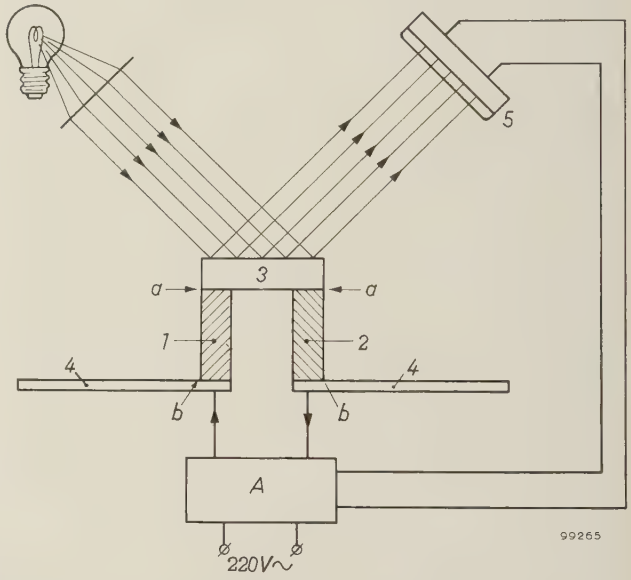


Fig. 1. Schematic representation of apparatus for automatic dew-point determination. Components *1*, *2*, *3* and *4* together constitute the whole Peltier element. *1* and *2* are two different semiconductors. *3* silver plate with reflecting surface. *4* copper plates acting as cooling fins for the warm junctions; they also serve as electrical terminals. When a current flows, heat is absorbed at the junctions *a* and developed at the junctions *b*. As the mirror mists over, less light falls on the photoresistor *5*, and the circuit *A* then feeds a smaller current through the Peltier element. In the state of equilibrium the temperature of *3* is equal to the dew point.

the current to the blocks. The junctions *b* are thus always at about room temperature. The upper surface of the silver plate is polished so as to produce a plane mirror. A parallel beam of light, obtained from an electric bulb and a lens, is directed on to this mirror at an angle of approximately 45°. The reflected beam impinges on a small photoresistor of cadmium sulphide [2]). When the apparatus is

*) Zentrallaboratorium Allgemeine Deutsche Philips Industrie G.m.b.H., Aachen Laboratory.

[1]) For an introduction to the various thermoelectric phenomena, and a recent review of developments, see A. F. Ioffe, Semiconductor thermoelements and thermoelectric cooling, Infosearch, London 1957.

[2]) See N. A. de Gier, W. van Gool and J. G. van Santen, Photoresistors made of compressed and sintered cadmium sulphide, Philips tech. Rev. **20**, 277-287, 1958/59 (No. 10). The resistor used here is of the type shown on the right in fig. 11 of that article.

switched on, so that current flows through the blocks and the reflector in the right direction, heat is developed at the junctions *b* and withdrawn from the junctions *a*. (Just as much heat is withdrawn from the junctions *a* as if the blocks *1* and *2* were directly connected with each other; the same applies to *b*.) As soon as the temperature of the silver plate drops below the dew point, the surface of the strip mists over and ceases to reflect specularly. The incident light is now reflected diffusely and less light reaches the photoresistor. As a result, the current flowing through this resistor also decreases. By means of an appropriate circuit, represented in the figure by *A*, the photocurrent controls the current flowing through the Peltier element in such a way that the cooling diminishes when the reflector mists over. The entire system thus constitutes a closed control loop. The temperature of the silver plate then rises once more above the dew point, the mist disappears and the process repeats itself. Effective damping of the control loop is ensured by suitably selecting the time constants, so that the equilibrium temperature is reached after only one or two oscillations; this temperature is the dew point.

It should be added that the diffuseness of reflection of the misted reflector depends on the "thickness" of the layer of condensate. This thickness, which we can express as the quantity of water per cm^2 of reflector surface, is proportional to the difference between the temperature of the reflector and the dew point and proportional to the time. After the temperature has dropped below the dew point, the nett cooling decreases but does not become negative *immediately:* the control is not an on-off but a proportional control, and the "integration" of dew on the mirror gives the system a measure of integral action [3]).

The fact that the equilibrium temperature coincides with the dew point and is not below it (it could not possibly be above it, because the reflector is then not misted and full cooling takes place) can be understood as follows. If the temperature of the reflector, and hence of the water condensed on its surface, is below the dew point, the mist will become increasingly dense and the cooling will decrease. Only when this temperature rises to the dew point can the mist achieve a steady state and the cooling remain constant.

[3]) See for example H. J. Roosdorp, On the regulation of industrial processes, Philips tech. Rev. **12**, 221-227, 1950/51. A distinctive feature of integral action in control systems is that the difference between the desired value and the value which the controlled quantity actually assumes in the state of equilibrium — the offset — is necessarily equal to zero.

The error in the dew-point determination with a hygrometer of this type can easily be made less than $\frac{1}{2}$ °C.

A working form of the circuit *A* (fig. 1) is illustrated in *fig. 2*. Since the resistance of the Peltier element *P* is low but the current through it fairly high (as much as 7 A), the simplest method of obtaining this current is to use an A.C. mains supply

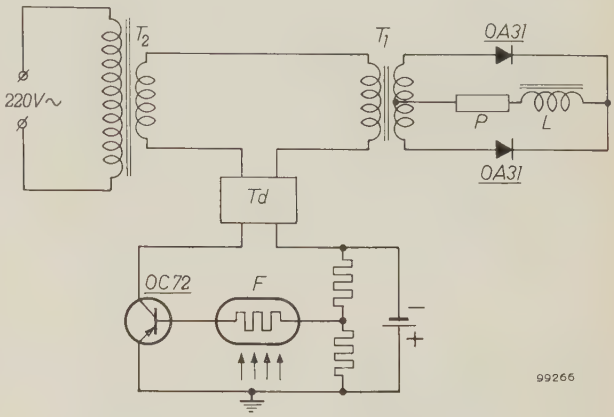


Fig. 2. Simplified diagram of the circuit. *P* Peltier element. *L* choke. T_1 and T_2 transformers. *Td* transductor (saturable reactor). *F* photoresistor, type B 873103. There are also two OA 31 germanium diodes and a type OC 72 transistor.

transformed down to a low voltage and then rectified. The primary of a transformer T_1 is connected via a transductor (saturable reactor) *Td* to the secondary of the mains transformer T_2. The direct current is obtained by means of two germanium diodes, type OA 31, and is smoothed by a choke *L* (capacitors are not so suitable for this purpose). The control current flowing through the saturable reactor is varied by a photoresistor *F* (type B 873103) with the aid of a transistor type OC 72.

In the side of the silver plate are a number of 1 mm holes into which a thermocouple or thermistor is introduced for measuring the temperature.

A thermocouple made of the same materials as the Peltier element is particularly suitable. Such a thermocouple gives a very high e.m.f. (340 μV/°C) and has a low electrical resistance, so that the whole measuring circuit can easily be given a constant resistance; in this way it is possible to measure the current in the circuit instead of the e.m.f. The current may easily be as high as 10 mA.

A complication, however, is that a milliammeter having an internal resistance of less than 1 ohm is required, and moreover this resistance must not be temperature-dependent.

Fig. 3a shows a photograph of the complete apparatus, and fig. 3*b* the Peltier element with the copper plates acting as cooling fins. These plates stand vertically in the instrument, close to the rear

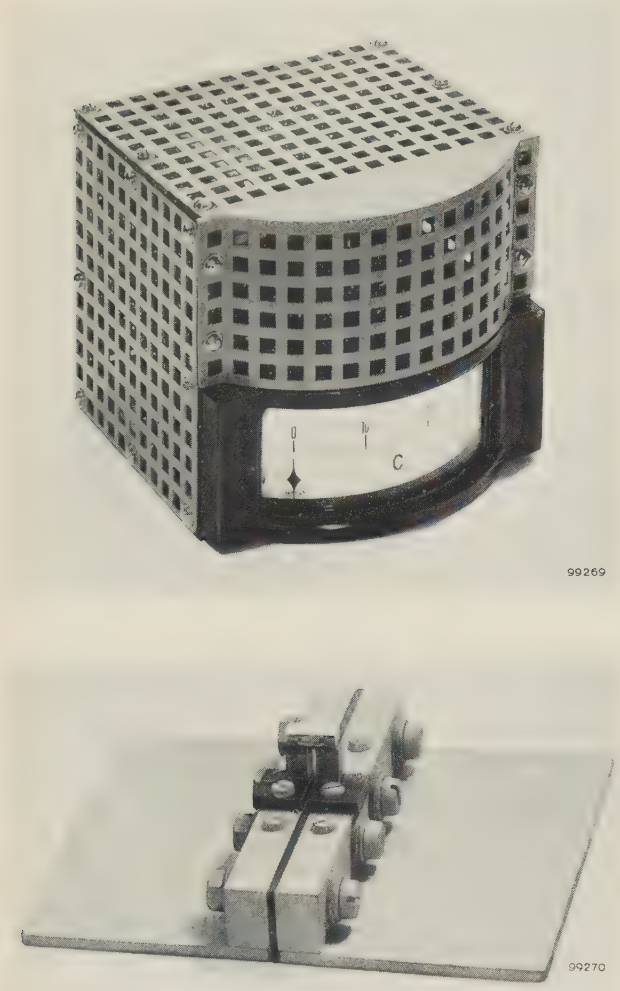


Fig. 3. *a*) Automatic dew-point hygrometer using Peltier cooling. *b*) Peltier element with copper plates acting as cooling fins. In the instrument these plates are vertically positioned close to the rear wall.

wall. The space above the millivoltmeter contains the lens system and the electrical circuit. The millivoltmeter is calibrated in °C.

We shall now consider at greater length the process of cooling by means of the Peltier effect. First, it should be noted that there is a limit to the temperature depression attainable with a Peltier element. True, the temperature depression produced is proportional to the current flowing through the element, but on the other hand ohmic heat is also generated, which increases in proportion to the *square* of the current. There is thus always a certain current at which the ohmic heating exactly compensates the Peltier cooling. Furthermore, there is also heat transfer from the warm to the cold junction via the Peltier elements themselves.

In order to draw up the heat balance from which we can calculate the maximum attainable reduction in the temperature of the reflector, it is necessary to make some simplifying assumptions:

a) The thermal conductivity and the electrical conductivity of both the materials constituting the element are equal (with the semiconductors employed this is in fact virtually true).

b) The cooling fins are so effective that the "warm" junction is at room temperature; the rest of the element is thus colder than the surroundings.

c) The quantity of heat absorbed from the surroundings is negligible.

d) The electrical and thermal resistance of the silver plate (*3* in fig. 1) and of the copper plates (*4* in fig. 1) are negligible.

e) Half of the ohmic heating is effectively dissipated at the junctions *a* and the other half at the junctions *b*.

f) All material constants are independent of temperature.

For the contributions to the power balance that refer to the cold junction (i.e. at the silver plate) we can now write the following simple formulae:

Peltier heat withdrawn: $W_1 = PI$.
Ohmic heat developed: $W_2 = \frac{1}{2} I^2R$.
Heat from the warm junction flowing per unit time through the Peltier elements: $W_3 = L\Delta T$.

(R is the electrical resistance of blocks *1* and *2* in series, and L is the thermal conductance of the blocks *1* and *2* in parallel.)

In the state of equilibrium W_1 must be equal to the sum of W_2 and W_3. Hence:

$$L\Delta T = PI - \tfrac{1}{2} I^2R. \quad . \quad . \quad . \quad . \quad . \quad (1)$$

By differentiating the right-hand side with respect to I and equating the result to zero, we find the current at which ΔT reaches its maximum value: $P - IR = 0$, whence $I_{max} = P/R$. From (1) we deduce that this value of ΔT is equal to:

$$\Delta T_{max} = \tfrac{1}{2} P^2/RL. \quad . \quad . \quad . \quad . \quad (2)$$

If, as is done in the dew-point hygrometer, the Peltier elements have the form of blocks of cross-section F and length l, it is a simple matter to express R and L in terms of the dimensions of the blocks. We find: $R = 2l\varrho/F$ and $L = 2F\lambda/l$, where ϱ is the resistivity and λ is the thermal conductivity of the materials used. In this case (2) becomes:

$$\Delta T_{max} = P^2/8\varrho\lambda. \quad . \quad . \quad . \quad . \quad . \quad (3)$$

It should be noted that the dimensions of the Peltier elements no longer occur in (3). Indeed, it can be shown that, whatever the form of the wires,

their dimensions are immaterial. If, for example, we try to reduce the ohmic heat by using short and thick blocks, we find that the "cold" gain is exactly offset by the increased flow of heat from the warm to the cold junction.

In cases where the above simplifying assumptions are not permissible, the foregoing is of course no longer true. The maximum temperature difference obtainable is then smaller than follows from (2) or (3). An important practical instance is where the Peltier element is required to supply a finite cooling *power*. Assumption *c*) is then no longer valid and a fourth term must be included in equation (1).

Combinations of substances for which the product $P^2/\varrho\lambda$ is relatively very large are to be found amongst the semiconducting compounds. A particularly large value of P can be obtained by combining an n-conductor with a p-conductor [4]). Of the two materials used in our case, one consists of 80% (by weight) Bi_2Te_3 + 20% Bi_2Se_3+0.02%AgI (n-conductor) and the other consists of 60% Sb_2Te_3 + 40% Bi_2Te_3 + 0.05% Ag. In our dew-point hygrometer they are used, as we have seen, in the form of blocks, 9 mm in length and of breadth and thickness 5 mm. At 3 A this element gives a temperature drop of about 30 °C. The maximum temperature drop is about 45 °C, which is reached when I is approximately 7 A [5]).

The above-described hygrometer using Peltier cooling is a variant of an automatic apparatus previously developed in our laboratory, which was based on the same principle (a change from specular to diffuse reflection caused by the misting of a reflector) but in which the cooling was effected by means of liquid air. It will perhaps be useful to give a brief description of this apparatus.

In this case the temperature of the silver reflector is automatically regulated by combining constant cooling with variable heating. The constant cooling is achieved in the following way. From a Dewar vessel filled with liquid air there protrudes one end of a bent rod of copper, about 10 mm in diameter. To the upper end is attached a disc of a material which is a poor thermal conductor (thermal resistance R_w), followed by a polished silver disc (*fig. 4*). The part of the rod projecting from the Dewar vessel is surrounded by thermal insulation. Because of its high thermal conductivity the entire rod assumes practically the temperature of liquid air (−194 °C). At the commonly occurring dew points (0-25 °C) the temperature difference between rod and reflector is thus always approximately the same, viz. about 200 °C: hence a practically constant heat flux passes from the silver disc to the copper rod. The thickness of the disc of insulating material is such that

$$200\ °C/R_w = 2\ \text{watts}.$$

The (variable) heating is provided by means of a turn of resistance wire wound around the rim of the silver disc and electrically insulated from it by mica. Thermal insulation is provided on the outside. The heater wire is connected via an ordinary loudspeaker transformer to a power amplifier valve which is driven by a 50 c/s alternating voltage. By varying the amplitude of this alternating voltage the power supplied to the disc can be varied between 0 and 4 W. With constant cooling at 2 W the total power is thus variable between −2 and +2 W. (When the disc is cold, the heat it receives from the ambient air is of the order of some tenths of a watt, which is by comparison insignificant.)

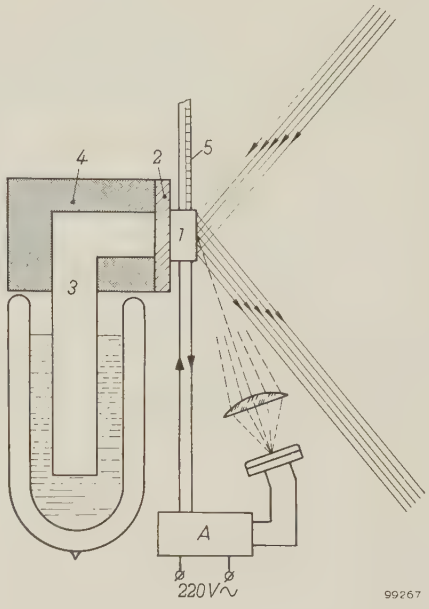


Fig. 4. Principle of dew-point hygrometer using liquid air. Constant cooling of the silver plate *1* is achieved by connecting it via a thermal resistance *2* to a copper rod *3*, the other end of which is immersed in a Dewar vessel containing liquid air; *4* thermal insulation, *5* mercury thermometer (may be replaced by a resistance thermometer or thermocouple). When the surface of *1* is clear, the beam of light bypasses the lens. When the surface mists over, some light is diffusely reflected on to the photocell connected to *A*; the latter then sends a current through the heater wire wound around *1*.

The automatic adjustment to the dew point is arranged in this case such that specularly reflected light from the clear surface of the silver disc just bypasses a lens in whose focal plane the light detector is situated. When the reflector becomes misty, the light is diffusely scattered. As long as the film of mist is still thin, however, most of the light is scattered in directions that make a fairly *small* angle with the direction of the reflected beam, so that quite a lot of light reaches the lens and hence the photocell or photoresistor. The relation between the photocurrent and the thickness of the film of mist is represented in *fig. 5*. At small thickness the photocurrent is proportional to the thickness of the film; it is in this region that the apparatus functions. As soon as light impinges on the photocell, the photocurrent gives rise to a voltage which drives the power valve referred to above, in such a way that the current through the heater wire (the anode current of the valve) increases as the mist thickens.

The characteristic difference between this instrument and the other is that in this case the current that regulates the temperature of the silver disc (the current through the heating

[4]) Concerning n and p conductors, see e.g. J. C. van Vessem, The theory and construction of germanium diodes, Philips tech. Rev. **16**, 213-224, 1954/55.

[5]) For comparison it may be noted that the combination of metallic conductors having the highest Peltier coefficient (Bi with Sb) can give a maximum temperature drop of just over 10 °C.

wire) must *increase* in conditions under which the current through the Peltier element should *decrease*, and vice versa. This is achieved by causing the specularly reflected light to bypass the light detector: the latter receives light only when the reflector is misted and thus reflects diffusely, the more so the thicker the film of mist. Here, then, the amount of light

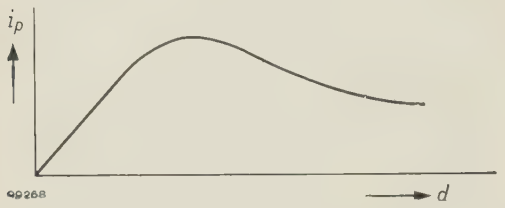


Fig. 5. Relation between the photocurrent i_p and the "thickness" d of the mist on the mirror. At small values of d the relation is linear.

incident on the detector varies from *zero* upwards, whereas in the apparatus using Peltier cooling it varies from a large to a smaller value. The former way of working may often be advantageous but is not used in the new hygrometer; the high sensitivity of the CdS photoresistor makes the other method no less satisfactory.

Finally it should be noted that the current through the resistance wire need not necessarily be a direct current, as required for Peltier cooling. It is thus possible to use a simple A.C. amplifier with transformer coupling to the output load; the latter can therefore be a short resistance wire through which a fairly large current flows. At the input of the amplifier, an alternating voltage is applied across the photocell.

In conclusion it may be said that the instrument which is the subject of this article is a nice example of the multifarious applications of semiconductors. They are used here (cf. fig. 2) not only in the Peltier element but also in the photoresistor, the transistors, the germanium diodes and finally in a resistance thermometer employing a thermistor.

Summary. The most accurate method of measuring the humidity of a gas is to determine the dew point. The cooling required for this purpose can be obtained by utilizing the Peltier effect of a combination of certain semiconductors. A description is given of a dew-point hygrometer based on this principle (it is a variant of an instrument using liquid-air cooling, also described). In this apparatus a silver reflector (5 mm thick, surface area 75 mm^2) is cooled to below the dew point. A beam of light is directed from this reflector on to a cadmium-sulphide photoresistor. When the reflector mists over, the beam is diffusely reflected; the photoresistor therefore receives less light and the photocurrent decreases. The latter controls the current through the Peltier element. The control loop thus formed is in equilibrium when the temperature of the reflector is equal to the dew point. The error in the measurement is less than $\frac{1}{2}$ °C.

A TRANSISTORIZED RADIATION MONITOR

by M. van TOL and F. BREGMAN.

621.317.794:621.375.4

This article adds the first measuring instrument to the long list of transistor-equipped devices that have been described in this Review in recent years. The emphasis here lies on the special requirements the circuit of a measuring instrument has to satisfy.

Radiation monitors are simple portable instruments for the detection and measurement of ionizing radiations. They are used in laboratories and hospitals whose staffs work with radioactive substances or with apparatus generating ionizing radiations (X-ray equipment, for example). Another field of application is the tracing of radioactive ores. These simple radiation meters are also suitable for special purposes such as in portable instruments for measuring the level of liquefied gases in cylinders [1]).

Batteries are the obvious form of power supply for portable radiation monitors, and since it is important that their dimensions and weight be small, the designer will be concerned to keep current consumption as small as possible. A radiation monitor which contained only one electronic tube apart from its Geiger-Müller tube, and which in consequence was very economical in operation, was described in this Review in 1953 [2]). It nevertheless required three batteries — a 40 volt anode battery, a 1.5 volt heater battery and a battery for giving the control grid a negative bias of 14 volts. The latter was rendered necessary by the special circuit developed to operate with only one tube.

Small portable electronic devices represent an extremely worthwhile field of application for transistors. In such devices the fullest advantage is taken of the characteristic properties of transistors, namely the small bulk, the low power consumption (and hence low heat dissipation) that results from the lack of heater currents, and the modest voltage requirements that can be met by small batteries, i.e. batteries comprising few cells. A prototype two-transistor radiation monitor, supplied from a single three-volt battery, was in fact developed in the Eindhoven Research Laboratories as early as 1954.

Transistors have less convenient aspects too, these being the spread in characteristics displayed by different individuals of the same type and the sensitiveness of their characteristics to temperature changes. Forming as they do an obstacle to precision, these drawbacks are ordinarily more serious in a measuring instrument than in an audio-amplifier, for example, where no great harm is done if the amplification changes slightly. However, a radiation monitor is a measuring instrument in which the demerits of transistors are not felt so keenly. For one thing, the main requirements are reliability and sensitivity rather than high accuracy; what is more, the measurement in question is based on pulse-counting, and for this purpose one can employ circuits in which the transistor characteristics have little or no influence on the actual reading. It is understandable that reliability should be given first place, for the purpose of the monitor is to measure radiation that is hazardous to human beings. Great sensitivity is called for because in certain cases it may have to detect radiation levels well below the human tolerance level. These requirements can certainly be satisfied just as well with transistors as with tubes.

Two commercial versions, types PW 4014 and PW 4012, have finally been evolved from the prototype. Equipped with four and five transistors respectively, the commercial models (illustrated in *fig. 1*) do not differ in any essential respect from the prototype, merely being improvements thereon.

We shall now discuss the circuit of the prototype, given in *fig. 2*. It can be divided into the following parts:

(*a*) H.T. supply circuit.

(*b*) Geiger-Müller tube supplied from this circuit, and delivering electrical pulses when exposed to ionizing radiation.

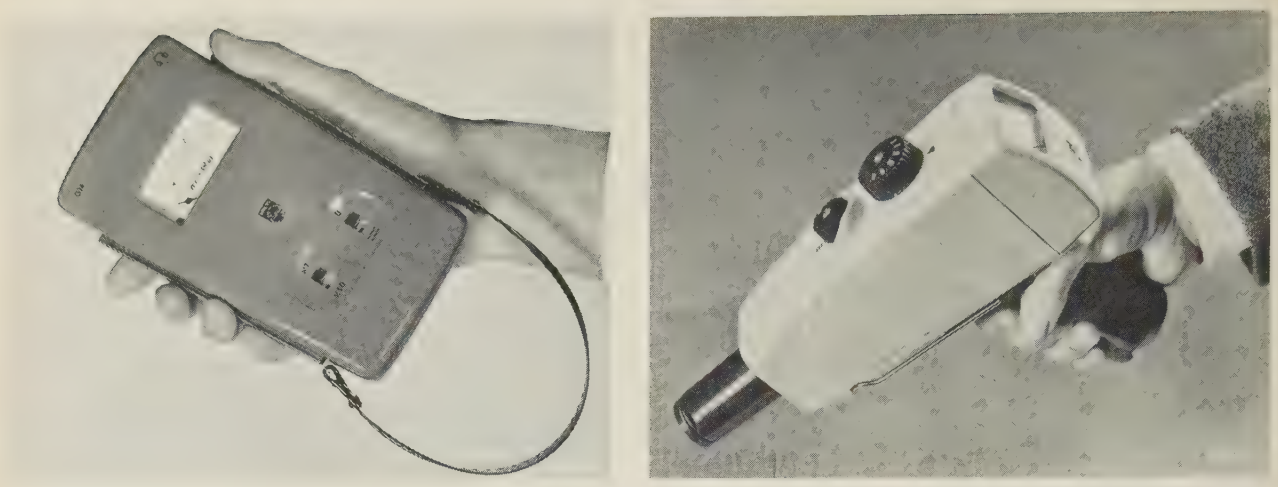
(*c*) Pulse-shaper that standardizes the height of the pulses.

(*d*) Diode pump-circuit which, supplied with a random sequence of pulses of uniform height, transforms them into a direct current whose value is proportional to the average number of pulses per second. This current is measured with the aid of meter *M*.

The H.T. supply is generated by a simple oscil-

[1]) See for example J. J. Arlman and H. N. L. Hoevenaar, Een niveaudetector voor de praktijk, De Ingenieur **71**, Ch. 8 - Ch. 9, 1959 (No. 19) (in Dutch).

[2]) G. Hepp, A battery-operated Geiger-Müller counter, Philips tech. Rev. **14**, 369-376, 1952/53.



a *b*

Fig. 1. Two commercial versions of a transistorized monitor for hard γ and β radiation which have been evolved from the prototype developed in the Philips Research Laboratories in 1954 and discussed in the present article.
a) Type PW 4014, pocket radiation monitor.
b) Type PW 4012, pistol model.

lator operating at a frequency of about 10 kc/s, using a transistor T_1, type OC 71. The voltage delivered by the oscillator is stepped up in the transformer and applied to a cascade circuit of three stages, which converts the alternating voltage into a direct voltage of treble its amplitude. The cascade circuit is equipped with selenium diodes.

A direct voltage of the same value could be obtained from a cascade circuit having some other number of stages and a transformer with a different output voltage. Losses in the cascade circuit increase with the number of stages; those in the transformer increase with the transformation ratio. The overall loss curve exhibits a shallow minimum corresponding to the use of three stages, as in the circuit under discussion; however, the particular transformation ratio and number of cascade stages decided upon make hardly any differences to the bulk and weight of the circuit.

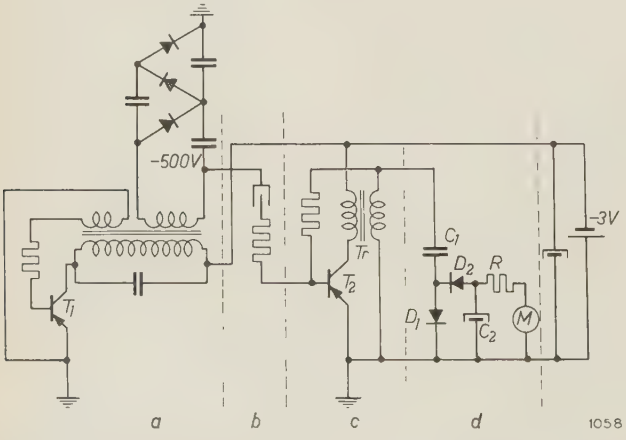


Fig. 2. Basic circuit of the prototype two-transistor radiation monitor. *a*: high-tension supply circuit for Geiger-Müller tube. *b*: G.M. tube with series resistor. *c*: pulse-shaper. *d*: diode pump-circuit with milliammeter *M*. The circuits of the commercial versions appearing in fig. 1 do not differ in essentials from that shown here.

The G.M. tube is of type 18 503 or 18 504; in both types the quenching gas is a halogen. On account of their robustness and almost unlimited life, these types of tube are particularly suitable for use in an instrument whose prime requirement is dependability. Moreover, their characteristics exhibit a long "plateau" with a particularly gentle slope (*fig. 3*). The high tension across the tube can therefore vary within wide limits (370 to 650 V) without greatly affecting the rate at which the tube delivers pulses when exposed to radiation of constant intensity [3]). For this reason it is possible to use a high-tension supply circuit of such simple design; for the same reason, there is no need to pay overmuch attention to the characteristics of transistor T_1. Provided the oscillator oscillates — and that is not asking a great deal — a usable high-tension supply will be available.

The voltage supplied is about 500 V, and the average current taken by the G.M. tube does not exceed 30 μA. The maximum power required is therefore 15 mW. A three-volt battery supplies the oscillator (and the other circuits). The high-tension circuit has an efficiency of about 60% and therefore takes less than 0.01 A from the battery — a very low current for a rod battery.

[3]) K. van Duuren, A. J. M. Jaspers and J. Hermsen, G. M. counters, Nucleonics **17**, No. 6, 86-94, 1959.

As already stated, the value of the high-tension voltage has little effect on the *rate* at which the G.M. tube delivers pulses. It does, however, have an effect on the *height* of the pulses. Since there is no stabilization of the H.T. supply, the average current through the G.M. tube cannot be used directly as a measure of the count rate. For this reason the pulses are fed to the pulse-shaping circuit *c* (see fig. 2), which consists of a type OC 71 transistor T_2 and a transformer *Tr*. Circuit *c* is a blocking oscillator [4]). T_2 acts as a switch that closes each time the G.M. tube delivers a pulse. When this happens, the primary winding of *Tr* carries the full battery voltage (3 V). The secondary winding of *Tr* thus has induced in it a voltage pulse whose amplitude is related to the battery voltage by the ratio of the numbers of turns. The time that T_2 remains in the conductive state, and hence also the width of the voltage pulse in the secondary winding of *Tr*, depend

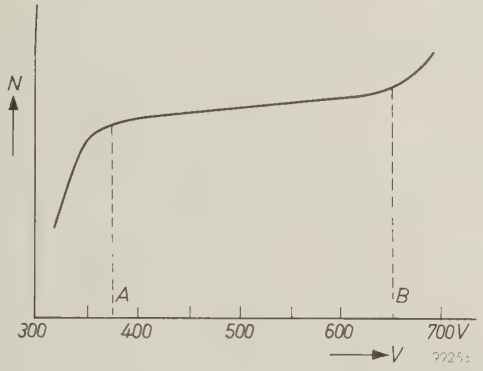


Fig. 3. Count-rate characteristic of a type 18 504 (halogen-quenched) Geiger-Müller tube. *N*, the number of pulses the tube delivers per second when exposed to a constant radiation intensity, is plotted as a function of the voltage *V* across the tube. The characteristic has a particularly low plateau slope (the "plateau" is the part of the curve between *A* and *B*).

amongst other things on the characteristics of the transistor. These also have an influence on the slope of the pulse edges, but the height of the pulses depends on the battery voltage alone.

A diode pump-circuit *d* measures the rate at which the voltage pulses, now of standard height, are produced. The circuit owes its name to the analogy with a piston force pump, in particular as used for compressing a gas. Diodes D_1 and D_2 represent the suction and delivery valves respectively; storage capacitor C_1 may be compared with the barrel of the pump, and the much larger buffer capacitor C_2 with the tank into which gas (i.e. electrons) is being pumped. However, charge continually drains away from the latter via resistor *R* and ammeter *M*.

A characteristic property of the circuit is that the current through *M* is dependent on the number and the height of the pulses entering it but not — provided they are sufficiently wide, allowing C_1 to be fully charged — on their duration or the slope of their edges. (Compare with the piston force pump, which displaces a quantity of gas that is dependent on the number of strokes per sec and on the pressure of the gas entering the pump, but not on the manner in which the barrel fills up, provided only that it fills with gas up to the full inlet pressure.) The diode pump-circuit is therefore very suitable for handling the pulses of uniform height that are delivered by the pulse-shaper *c*. It is possible, by a correct choice of values for C_1, C_2 and *R*, to make the current through *M* proportional, or nearly so, to the pulse rate. A recent article in this Review [5]) may be referred to for a more detailed discussion of the way in which a diode pump-circuit functions.

We have seen above that the constituent parts of the monitor have been adapted to one another with the object of getting a meter reading that is independent of transistor characteristics. We may recapitulate as follows.

The H.T. voltage yielded by circuit *a* (fig. 2) depends on T_1. It is applied to G.M. tube *b*, which delivers pulses at a rate which is independent of the voltage applied and hence of T_1. The height of the pulses *is* dependent on T_1. Pulse-shaper *c* gives these pulses a constant *height*. Their *shape* (i.e. their duration and the slope of their edges) depends, however, on the characteristics of T_2; the influence of T_1 has been eliminated. Diode pump-circuit *d* converts these pulses of standard *height* into a current that is independent of their *shape*. Thus the influence of T_2 has also been eliminated.

Even so, the current through the measuring instrument is still dependent on the battery voltage, for the height of the pulses delivered by circuit *c* is proportional to that voltage (see above). Facilities for checking the battery voltage have therefore been provided in the commercial versions in fig. 1, the indication being given on the meter (*M*).

The ranges measured by pocket radiation monitor PW 4014 are 0 to 3 mr/h (milliröntgen per hour) and 0 to 30 mr/h [6]). The pistol model, PW 4012, has the ranges 0 to 1 mr/h, 0 to 10 mr/h and 0 to 100 mr/h. The design of both models allows for the

[4]) For the functioning of a blocking oscillator see, for example, B. Chance *et al.*, Waveforms, No. 19 of the Radiation Laboratory Series, McGraw-Hill, New York 1949.

[5]) J. J. van Zolingen, Philips tech. Rev. **21**, 134-144, 1959/60 (No. 4/5), in particular p. 138 et seq.

[6]) Röntgen, röntgen per hour and other units relating to ionizing radiation were discussed in N. Warmoltz and P. P. M. Schampers, A pocket dosemeter with built-in charger, for X-radiation and gamma radiation, Philips tech. Rev. **16**, 134-139, 1954/55.

connection of a type 18 509 counter tube, which extends the PW 4014 ranges to 0-60 and 0-600 mr/h, and the PW 4012 ranges to 0-20 and 0-200 mr/h. To give some idea of the significance of these figures, we would add that 5 röntgens per year is regarded as an acceptable dose for persons exposed to radiation by reason of their occupation. In a working year of 2000 hours this amounts to 2.5 mr/h. The tolerance dose for other classes of the population is lower.

A circuit giving a new kind of meter scale

A new circuit has meanwhile been developed in the laboratory in which the scale of the meter is progressively compressed (though *not* logarithmically) at higher pulse rates. This means that, as for a logarithmic scale, more economical use is made of the available scale length than when a linear scale is used. The deflections obtained with this circuit differ from those of a logarithmic scale in that zero deflection corresponds to absence of radiation and full-scale deflection to an infinitely high radiation level. Consequently the needle never goes beyond the extremes of the scale — a very convenient feature in a radiation monitor: the user always has immediate indication of the radiation level without range-switching.

The principle of the new circuit is as follows. *Fig. 4* shows a flip-flop circuit with two transistors T_3 and T_4. Negative pulses from the G.M. tube arrive at terminal A at an average rate of N per second. After such a negative pulse on A the circuit is left in the stable state in which T_3 conducts; suppose

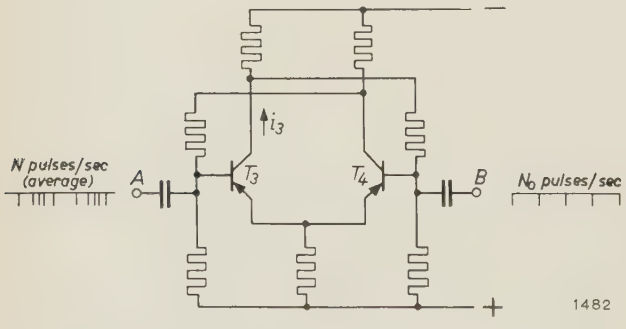


Fig. 4. Flip-flop circuit in which terminal A goes to the G.M. tube, and terminal B to a pulse generator that delivers pulses at a constant rate of N_0 per sec. The mean current through transistor T_3 is a measure for N, the average number of pulses delivered by the G.M. tube per sec.

that the current i_3 then flowing through T_3 is I_3. A pulse generator supplying negative pulses of constant repetition frequency N_0 is connected to terminal B. After each of these pulses the circuit is left in the state in which T_3 is non-conducting. The mean current through T_3 is now a measure for N; this can be seen with the help of *figs. 5a* and *b*. In each diagram the pulses arriving at A, those

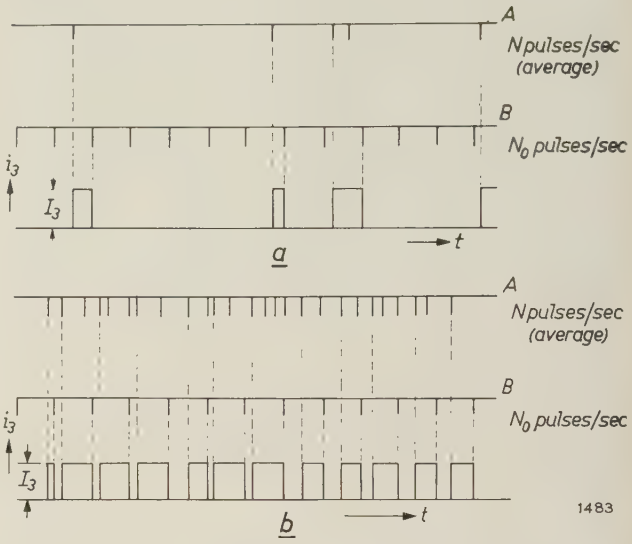


Fig. 5. Diagram to explain the functioning of the circuit in fig. 4.

arriving at B and the current through T_3 are plotted one below the other as functions of time; the current i_3 through T_3 can only have one of two values, zero or I_3. When T_3 has been rendered conducting by a pulse from the G.M. tube, a current I_3 flows until stopped by the next pulse of the regular train from the generator. It will be seen from fig. 5*a* that if N is much less than N_0, as in that diagram, then the mean current through T_3 will have a low value. Indeed, if the G.M. tube stops delivering pulses altogether, the mean current will be zero. Fig. 5*b* shows the situation when N greatly exceeds N_0. Here T_3 is conducting almost all the time. At very high radiation levels the mean current through the transistor has a value approaching I_3.

The two extreme cases $N \ll N_0$ and $N \gg N_0$ lend themselves to a simple quantitative treatment, as follows.

For the case $N \ll N_0$, almost every one of the N pulses delivered by the G.M. tube per sec will reach T_3 when it is in the cut-off state. At any instant the average time that will elapse before the next pulse arrives from the generator is $\frac{1}{2}(1/N_0)$ sec. Roughly, then, T_3 conducts for a proportion of the time given by $\frac{1}{2}N/N_0$, and the mean current through it is $\frac{1}{2}(N/N_0)I_3$. The lower end of the scale therefore has a linear character.

The case $N \gg N_0$ can be dealt with analogously. Of the N_0 pulses delivered by the generator each second, almost every one reaches T_3 when it is in

the conducting state. At any instant the average time required for the next pulse to arrive from the G.M. tube is $1/N$ sec. (The factor $\frac{1}{2}$, appropriate to pulses arriving in a *regular* train, is absent here on account of the statistical distribution of the pulses delivered by the G.M. tube.) Roughly, then, T_3 is cut off for a proportion of the time given by N_0/N, and the mean current through it is $(1-N_0/N)I_3$. At the upper end of the scale, therefore, the deflection of the needle approaches the full-scale value (I_3) hyperbolically.

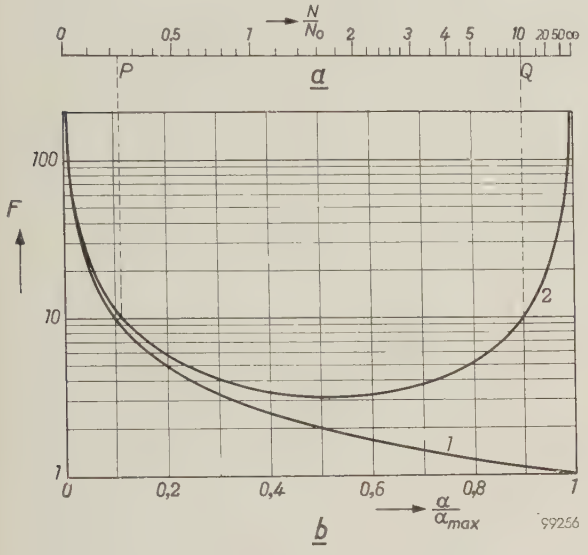


Fig. 6. *a*) Scale obtained with the new circuit; it provides readings of N/N_0, the ratio between the average number of pulses delivered by the G.M. tube per sec and the constant number of pulses delivered by the pulse generator per sec. *b*) The quantity F indicates the percentage measuring error arising from a (constant) deflection error of 1% of α_{max}, the full-scale angular deflection. F is plotted logarithmically as a function of α/α_{max} for a linear scale (curve *1*) and for the new scale (curve *2*).

The general formula for the mean current $\overline{i_3}$ through T_3 is:

$$\overline{i_3} = I_3\left(1 - \frac{1 - e^{-N/N_0}}{N/N_0}\right).$$

A scale in terms of the ratio N/N_0, based on the above formula, is drawn in *fig. 6a*. If N_0 is made equal to the G.M. pulse rate corresponding to 1 mr/h (this rate depends on the properties of the tube employed), then we obtain a scale whose unit division corresponds to 1 mr/h.

Comparison of the new scale with a linear scale makes clear how economical the former is in its use of the available scale length. The most obvious basis for comparing two meter scales, both of which are used for measuring a quantity x, is the relative error of measurement, $(\Delta x)/x$. This error in x proceeds from the error $\Delta\alpha$ in the deflection α undergone by the needle, and is given by

$$\frac{\Delta x}{x} = \frac{1}{x}\frac{dx}{d\alpha}\Delta\alpha.$$

Hence the relative error in x per unit deflection error is equal to $(1/x)(dx/d\alpha)$; this expression, plotted as a function of α, is therefore a suitable yardstick for comparing the merits of different scales. In fig. 6*b* the dimensionless quantity

$$F = \frac{\alpha_{max}}{x}\frac{dx}{d\alpha}$$

is plotted as a function of another dimensionless quantity α/α_{max} to give curves that are independent of the unit in which α is expressed. α_{max} stands for the full-scale deflection. F denotes the percentage measuring error proceeding from a deflection error $\Delta\alpha$ that is 1% of α_{max}. If it is stipulated that F shall nowhere exceed 10, then it will be clear from fig. 6*b* that the linear scale has a usable range extending from $\alpha/\alpha_{max} = 0.1$ to $\alpha/\alpha_{max} = 1$, while that of the new scale extends from P to Q, i.e. from $\alpha/\alpha_{max} \approx 0.11$ to $\alpha/\alpha_{max} \approx 0.9$. The usable portion of a linear scale covers a measuring range whose upper limit occurs at a value 10 times greater than the value read at its lower limit. On the new scale in fig. 6*a* points P and Q occur at values of approximately 0.233 and 10 respectively; the upper limit thus occurs at a value 10/0.233 times or a good 40 times greater than the value read at the lower limit. *Fig. 7a* shows two linear scales covering adjoining ranges; together they cover the range from 0.3 to 30 mr/h with a percentage measuring error F nowhere exceeding 10. Two scales of the new type covering adjoining ranges, so chosen that the lower useful limit of the top scale again lies at 0.3 mr/h, are reproduced in fig. 7*b*; with the same maximum F value of 10, these have an overall range going up to 500 mr/h. A reasonable estimate can moreover be made up to 2000 mr/h.

Calibration of the new scale is particularly easy. The flip-flop circuit is held first in one and then

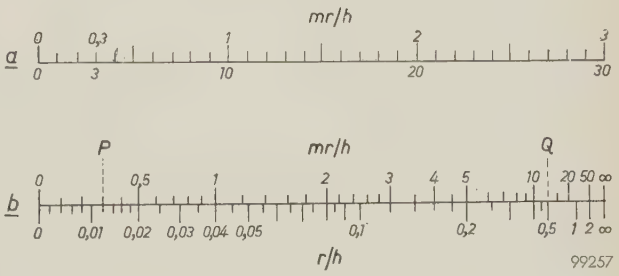


Fig. 7. *a*) Two linear scales which together cover the range from 0.3 to 30 mr/h with a percentage measuring error $F \leqslant 10$. *b*) Two complementary scales of the new type. Here the total range covered, over which $F \leqslant 10$, extends from 0.3 mr/h (as on the linear scale in *a*) to 500 mr/h (0.5 r/h).

in the other of its stable states; the circuit is adjusted so that the meter reads zero in the state in which T_3 is cut off and adjusted to give full-scale deflection in the state in which T_3 conducts. All that is now necessary is to give N_0 the desired value. The normal practice will be to place the monitor in the radiation field of a standard radioactive preparation and then to adjust the generator to give a pulse repetition frequency such that the correct meter reading is obtained. Thus only one point has to be calibrated on each measuring range, as in the case of a linear scale. The new scale is much more convenient to calibrate than a logarithmic scale. Owing to the lack of a zero, a logarithmic scale has to be calibrated at two points, and adjustment at one point entails readjustment at the other.

Summary. Portable radiation monitors are an obvious case in which transistors can be used to advantage. As measuring instruments, however, the circuit employed must be so designed that the readings are insensitive to the transistor characteristics (which may shift considerably with changes of temperature). A circuit of this kind is described. The authors also discuss a method of arriving at a (non-linear) scale that covers the whole range between zero radiation (zero deflection) and an infinitely high radiation level (full-scale deflection). Readings are sufficiently accurate (less than 10% error for a deflection error equivalent to 1% of the scale length) over a part of the scale extending from 0.1 to 0.9 of the scale length. The radiation-level reading at the upper limit of this range is a good 40 times higher than that at the lower limit. The instrument can easily be adjusted to bring any given radiation level within the usable range. The calibration procedure is particularly straightforward.

STRAY CAPACITANCES IN NEON INSTALLATIONS

by J. J. WILTING.

621.327.42

High-voltage neon-filled discharge tubes, used widely as luminous signs, can generate strong electrical transients. These are bound up with the stray capacitances in the installation. The article below deals with the favourable as well as the unfavourable consequences of these phenomena.

A gas discharge which is maintained by an alternating current of low frequency is periodically extinguished and must therefore be periodically reignited (twice in every cycle). The voltage needed for restarting the discharge depends on the nature and the pressure of the gas filling and on the geometry of the tube. In the case of long discharge tubes, as frequently used for luminous signs, the reignition voltage U_i is not much higher than the burning voltage U_b if the gas filling is mercury vapour (*fig. 1a*). With neon gas, on the other hand, U_i may well be three or four times higher than U_b (fig. 1*b*).

The high reignition voltage of neon tubes thus calls for a supply transformer with a particularly high no-load voltage, and is therefore quite expensive; on the other hand, an upper limit is set to the permissible transformer voltage by safety regulations. The presence of the stray capacitances of the transformer and of the high-tension cable connecting the tubes to the transformer is fortunate in this respect, in that it offers the possibility of generating a high voltage pulse at approximately the moment when reignition is needed. This facilitates the reignition and lowers the necessary no-load voltage of the transformer.

As the current periodically drops to zero (initiating the "dark period", i.e. a currentless interval lasting until the reignition), a transient is produced in the electrical circuit. This consists of a damped oscillating voltage superimposed on the voltage across the neon tubes, with a frequency of about 500 to 1200 c/s, which is thus substantially higher than the mains frequency. The sum of this damped oscillation (the above-mentioned voltage pulse) and the voltage of the temporarily unloaded transformer may be high enough for reigniting the neon tubes.

Capacitance (and stray capacitance is always present) is essential to the occurrence of this useful effect. The other side of the picture, however, is that the capacitance, after charging up to a high voltage, immediately discharges through the neon tubes after every reignition. The discharge takes the form of a strong current surge which makes the neon tubes burn unsteadily and shortens their life.

In the following we shall consider the means of promoting the generation of an effective voltage pulse whilst limiting the current surges to tolerable values.

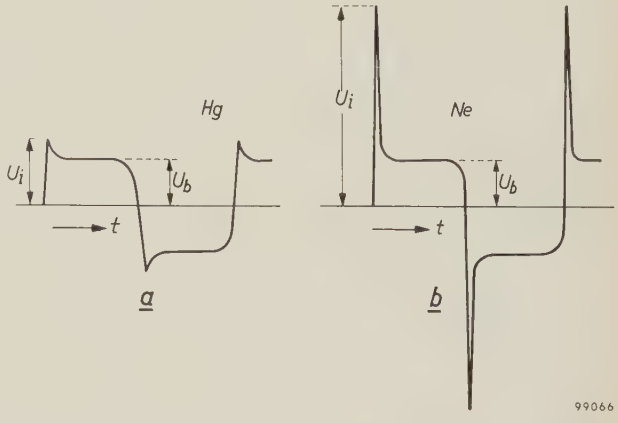


Fig. 1. *a*) In long discharge tubes containing mercury vapour (under low pressure) the reignition voltage U_i is not much higher than the burning voltage U_b.
b) In long neon tubes U_i is several times higher than U_b.

The voltage pulse for reignition

Transformers for neon tubes (one or more tubes in series) are generally designed with some magnetic-flux leakage, with the object of obtaining a specific internal reactance that will take up the difference between the no-load and burning voltages. Some designs in common use are shown in *fig. 2*. By means of magnetic shunts *Sh* the reactance is adjusted such that the current assumes its rated value. The advantage of the balanced designs (fig. 2*b* and *d*) over the unbalanced is that the secondary voltage may permissibly be twice as high as the maximum voltage with respect to earth as prescribed by safety regulations.

Since the tubes become conducting and non-conducting fairly abruptly, they may be considered, as regards their circuit behaviour, as a switch in series with (principally) a resistance: the switch opens at the moment $t = 0$ when the current drops to zero,

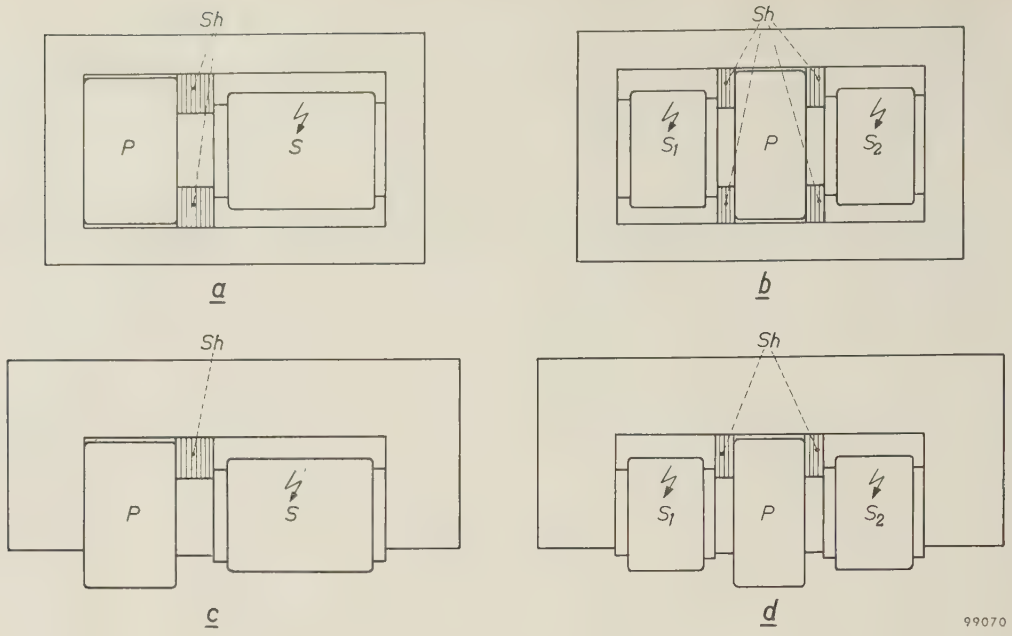


Fig. 2. Common designs of high-tension transformers for neon tubes. P primary coil. S secondary coil (in b and d divided into balanced halves, S_1 and S_2). a and b have cores of the shell type; c and d of the yoke type. Sh magnetic shunts for adjusting the leakage reactance to a specific value.

and closes again at the moment of reignition. We denote the no-load voltage by $e = E \sin(\omega t + \varphi)$; the voltage across the neon tubes at $t = 0$ (beginning of the dark period) is the superposition of the instantaneous value $e_0 = E \sin \varphi$ of the no-load voltage at $t = 0$ and the transient voltage Δe (*fig. 3*).

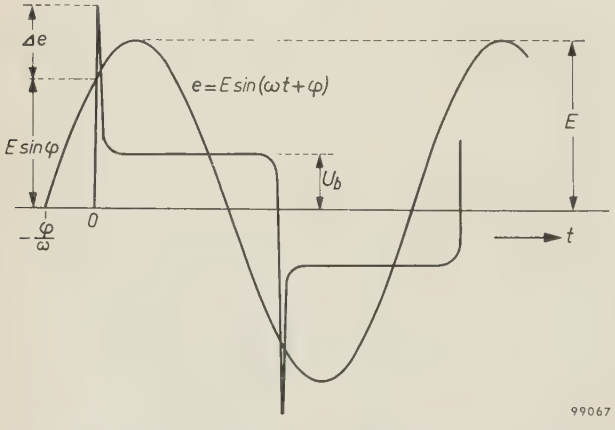


Fig. 3. Voltage waveform in a burning neon tube. No-load voltage of transformer: $e = E \sin(\omega t + \varphi)$. At the moment at which the current drops to zero (here also the moment of reignition) $t = 0$. The no-load voltage at that moment is $E \sin \varphi$; but superposed on this is a voltage Δe, supplied by the transient, which must be sufficient to ensure reignition. U_b = burning voltage.

The transient can be calculated to a good approximation on the basis of a simplified equivalent circuit. *Fig. 4a* shows the diagram of a balanced neon installation with the stray capacitances C_t of the transformer and C_k of the high-tension cable. Trans-

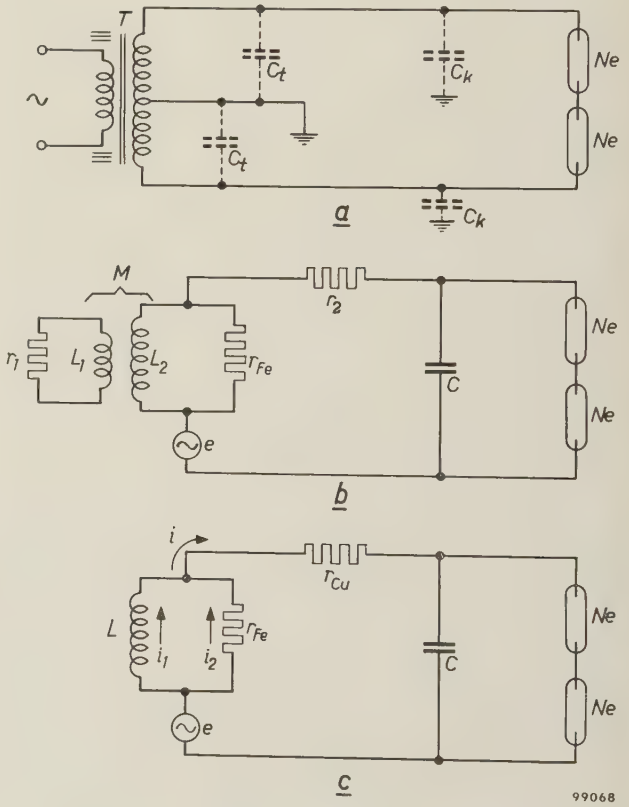


Fig. 4. *a*) Balanced transformer of the type in fig. 2*b* or *d*, supplying two neon tubes in series. C_t and C_k are stray capacitances of transformer and high-tension cable, respectively. *b*) Equivalent circuit of (*a*) with AC voltage source e placed in the secondary circuit. L_1 and r_1, L_2 and r_2 represent the inductance and resistance of the primary and secondary coils, respectively. M mutual inductance. r_{Fe} effective resistance due to iron losses. C total stray capacitance.
c) Simplified equivalent circuit. L leakage inductance. r_{Cu} effective resistance due to all copper losses.

ferring the voltage source to the secondary circuit and short-circuiting the primary coil produces the equivalent circuit in fig. 4*b*. All stray capacitances are now lumped together in the capacitance C. The resistance and inductance of the primary coil are, respectively, r_1 and L_1, and of the secondary coil r_2 and L_2. A resistance r_{Fe} is added to allow for the iron losses.

By further simplification, this circuit can be reduced to that of fig. 4*c*. Here the inductance L is given by

$$L = (1 - k^2)\, L_2,$$

where k is the coupling coefficient of the transformer. To a first approximation L and r_{Fe} may be regarded as constants.

The resistance r_{Cu} in fig. 4*c* allows for the total copper losses:

$$r_{Cu} = r_2 + \frac{L_2}{k^2 L_1} r_1.$$

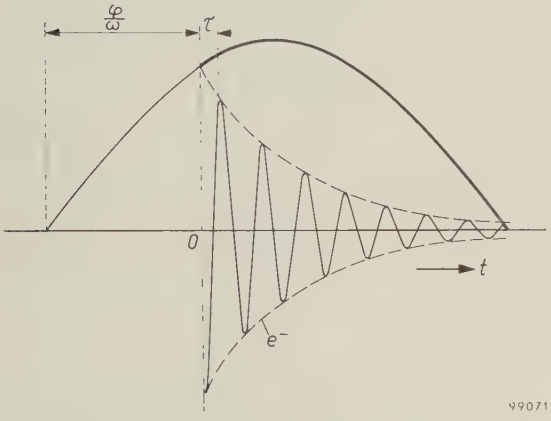


Fig. 5. The sine wave represents the no-load voltage, the damped oscillation is the transient.

Making certain minor approximations, we can derive for fig. 4*c* the following equation describing the variation of the capacitor voltage u with time t:

$$u(t) = E \sin(\omega t + \varphi) - E \sin\varphi \,.\, e^{-at} \cos \beta t. \quad (1)$$

In this expression ω is the angular frequency of the mains, a the damping constant and β the natural angular frequency of the excited "high-frequency" resonant circuit L-C; see *fig.* 5. For a and β we can write:

$$a \approx \frac{r_{Cu}}{2L} + \frac{1}{2 r_{Fe} C} \quad \ldots\ldots \quad (2)$$

and

$$\beta \approx \frac{1}{\sqrt{LC}}. \quad \ldots\ldots\ldots \quad (3)$$

An oscillogram of the transient is shown in *fig.* 6.

Equations (1), (2) and (3) are obtained as follows. Let the currents through r_{Cu}, L and r_{Fe} in fig. 4*c* be i, i_1 and i_2, respectively; then:

$$i = i_1 + i_2,$$

$$L \frac{di_2}{dt} = i_2 r_{Fe},$$

$$i_2 r_{Fe} + i r_{Cu} + u = e,$$

$$i = C \frac{du}{dt}.$$

Eliminating i, i_1 and i_2 from these expressions, we find a differential equation of the form

$$\frac{d^2u}{dt^2} + a \frac{du}{dt} + bu = ce + d \frac{de}{dt}, \quad \ldots \quad (4)$$

where a, b, c and d are certain functions of L, C, r_{Cu} and r_{Fe}. For the overshoot above the no-load voltage the periodic solutions of the differential equation (4) are the only ones of importance. These are of the form

$$u(t) = A \sin(\omega t + \gamma) + B\, e^{-at} \sin(\beta t + \delta). \quad . \quad (5)$$

This equation expresses the superposition of the no-load voltage of angular frequency ω and a damped oscillation of higher angular frequency β. The coefficient A and the phase angle γ follow from an easily found particular solution. If the initial conditions are known, B and δ can also be calculated.

The initial moment is the moment at which the current i reaches the zero value. The burning voltage then also falls to zero, so that we may write:

$$u(0) = 0.$$

From (5) it therefore follows that

$$A \sin\gamma + B \sin\delta = 0. \quad \ldots\ldots \quad (6)$$

Fig. 6. Oscillogram of transformer voltage e (no-load operation) and of the voltage u across a neon tube which passes current in one direction only, a rectifier being connected in series with it to make the oscillogram clearer. The transformer voltage alone, without the transient, is too low to guarantee reignition of the neon tube in every cycle. The superposition of the two curves in fig. 5 can be recognized in the waveform of u.

The second initial condition is given by the magnitude of du/dt at the moment $t = 0$. This depends on various conditions. Suppose

$$\left(\frac{du}{dt}\right)_{t=0} = U_0',$$

then from (5) we find:

$$U_0' = A\,\omega\cos\gamma + B\,\beta\cos\delta - B\,\alpha\sin\delta. \quad . \quad . \quad (7)$$

From (6) and (7) we now obtain:

$$\tan\delta = \frac{A\beta\sin\gamma}{A\,\omega\cos\gamma + A\,\alpha\sin\gamma - U_0'}.$$

Since the effect of the resistances r_{Cu} and r_{Fe} is negligible compared with that of the inductance L and the capacitance C, it follows from fig. 4c that, approximately,

$$A \approx \frac{E}{1-\omega^2 LC} \quad \text{and} \quad \gamma \approx \varphi.$$

Where the high-tension cables are not unduly long, $1/\omega C \gg \omega L$, hence $A \approx E$. With this approximation equation (5) becomes

$$u(t) \approx E\sin(\omega t+\varphi) - E\frac{\sin\varphi}{\sin\delta}\,e^{-\alpha t}\sin(\beta t+\delta). \quad . \quad (8)$$

Measurements have shown that, at all encountered values of U_0', the value of δ is only a few degrees smaller than 90°, so that $\sin\delta \approx 1$. Insight into the phenomena taking place can therefore be gained by putting the angle δ in (8) equal to $\frac{1}{2}\pi$

$$u(t) \approx E\sin(\omega t+\varphi) - E\sin\varphi\,.\,e^{-\alpha t}\cos\beta t,$$

and this is the above equation (1).

Finally, the damping factor α and the angular frequency β of the damped oscillation follow from (5) after writing A and B in terms of the constants a, b, c and d, and substituting for the latter in terms of the various circuit parameters:

$$\alpha = \frac{L + r_{Cu}\,r_{Fe}\,C}{2(r_{Cu}+r_{Fe})LC}$$

and

$$\beta = \sqrt{\frac{r_{Fe}}{(r_{Cu}+r_{Fe})LC} - \alpha^2}.$$

By approximation these expressions can be simplified to formulae (2) and (3) given above.

The required reignition voltage

The reignition voltage of a given neon tube depends principally on the time interval τ between extinction and reignition, and, where the tube is reignited by a voltage *pulse*, on the duration of this pulse. (The magnitude of the current through the tube before the extinction also plays some part, but this need not be considered here.)

The first of the two effects — the dependence on τ — is demonstrated in *fig. 7a* by curves *1* and *2*. The curves represent the reignition voltage as a function of τ measured on two neon tubes of the same diameter (13 mm), but measuring respectively 3.85 m and 1.85 m in length. The measurements were done with direct voltage which was switched off (electronically) for a variable time τ. Where τ is of the order of some microseconds the reignition voltage is approximately equal to the burning voltage. When τ is increased, the reignition voltage also increases until it reaches a maximum at $\tau \approx 1000$ μsec, where it is equal to the (initial) ignition voltage.

If the tube operates on alternating voltage, it is reignited by a short-lived peak and the situation is considerably less favourable. Because of the short duration of the pulse, the discharge can only be properly initiated if the peak voltage is much higher than the direct voltage that was sufficient for the measurements mentioned above. This situation is represented in fig. 7*b*. The lines *1* and *2* roughly correspond to curves *1* and *2* in fig. 7*a*, and thus

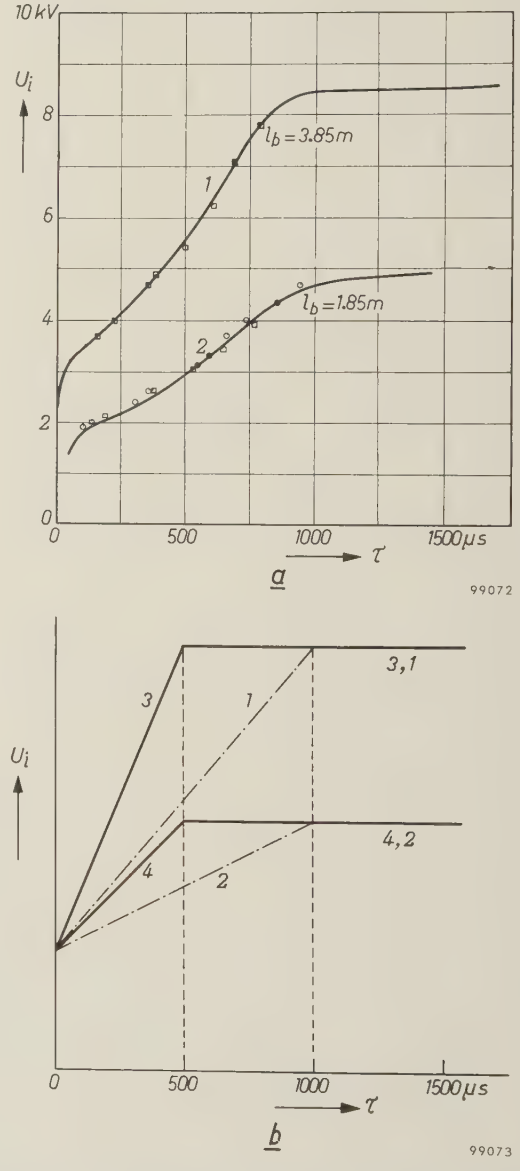


Fig. 7. *a*) Reignition voltage U_i for neon tubes (13 mm inside diameter) operated on DC, the supply being electronically switched off for a variable interval τ, as a function of τ. Curve *1* refers to a tube 3.85 m long, curve *2* to a tube 1.85 m long, both on a current of 25 mA.
b) Reignition voltage on AC operation (schematic); *3* refers to the 3.85 m tube, *4* to the 1.85 m tube. For comparison the DC lines *1* and *2* are shown, which roughly correspond to *1* and *2* in (*a*).

apply to direct voltage. In the case of alternating voltage having a waveform as shown in fig. 6 and a pulse duration of $\frac{1}{2}\tau$, the curves are as shown by *3* and *4* for neon tubes of lengths 3.85 and 1.85 m respectively: the reignition voltage has now risen to the ignition voltage at $\tau \approx 500$ μsec instead of at 1000 μsec.

In a normally burning neon tube the dark period is longest when the first positive peak of the damped oscillation only just reaches the reignition voltage. At the moment of reignition, a complete half-cycle of the damped oscillation has then elapsed since the moment $t = 0$ (fig. 5). As the oscillation has the angular frequency β, given by equation (3), the maximum dark period τ_{max} is:

$$\tau_{max} = \frac{\pi}{\beta} = \pi \sqrt{LC}.$$

In large neon installations (no-load voltage 8 kV) the values of L and C are generally such that $\tau_{max} > 500$ μsec. Neon tubes of 13 mm diameter (for which fig. 7 holds) therefore require the full ignition voltage for reignition in such an installation. At larger tube diameters the curves *3* and *4* in fig. 7*b* shift to the right; it may then well be, particularly in smaller installations with higher β, that the working point for reignition comes to lie on the sloping portion of the curve.

Relation of maximum tube length to transformer voltage

A factor of considerable practical importance in neon installations is the maximum length of tube that will operate efficiently on a given transformer voltage. This length is sharply limited. The longer the tube, the higher is the *required* reignition voltage, but the lower is the *available* reignition voltage. The truth of the latter statement will now be demonstrated below.

The first voltage peak $u(t)$ occurs at about the moment $t = \tau = \pi/\beta$. Substituting this value for t in equation (1), we obtain:

$$u(\tau) \approx E \sin\left(\frac{\omega}{\beta}\pi + \varphi\right) + E \sin\varphi \exp\left(-\frac{\alpha}{\beta}\pi\right). \quad (9)$$

If the tube length is changed, the current is restored to the rated value by adjusting the magnetic shunts. This is attended by a change in the phase angle φ. The reason is that an increase in tube length represents increased resistance in the circuit, and thus necessitates adjustment of the shunts to obtain a smaller leakage reactance. Both changes (larger resistance, smaller inductance) cause a reduction of φ, and hence a reduction of both terms on the right-hand side of equation (9) (the argument of the first term, $(\omega\pi/\beta) + \varphi$, is an angle in the *first* quadrant). These terms together constitute the available reignition voltage.

On the basis of (9) we shall now examine the measures that can be taken to keep the available reignition voltage adequate for as long as possible when increasing the length of the tube. The parameters that can be manipulated in this equation are the angular frequency β and the damping factor α. It is seen from (9) that the first term of the right-hand side increases when β is decreased, and that the second term increases when α/β is decreased. We thus require a reduction of β accompanied by a relatively greater reduction of α.

Let us first consider the possibilities of reducing the angular frequency β. This can be done by increasing C or L, or both. C is governed mainly by the length of the high-tension cable. The magnitude of C, hence effectively the length of the cable, is limited, however, because the surge currents increase with C; if they are excessive, they cause the tubes to burn unsteadily and shorten their life (see next section). The practical limit is reached with cables of a few metres long. As regards the inductance L, it should be noted that this quantity is not identical with the leakage inductance that determines the steady-state current. The rapid flux changes that give rise to the transient are superposed on the steady-state alternating field. The inductance L concerned with the transient is therefore determined by the incremental permeability μ_i of the type of laminated core steel employed; μ_i depends on the pre-magnetization due to the main magnetic field during the dark period. In designing the transformer one can therefore try to make L large (and hence β small) by using lamination steel having a high μ_i and by not saturating the core too strongly.

Let us now consider the ratio α/β, which should also be small. We have seen that the damping factor α is given by:

$$\alpha \approx \frac{r_{Cu}}{2L} + \frac{1}{2r_{Fe}C}. \quad \ldots\ldots \quad (2)$$

Inserting the numerical values encountered in practice we find that the second term, $1/(2r_{Fe}C)$, is predominant. At a given C, therefore, α is almost solely dependent on the iron losses. *Fig. 8* shows three curves of α plotted as a function of β for a transformer of the type 2×4000 V, 25 mA. As might be expected, α decreases when the iron losses are reduced by using either thinner laminations or a steel with a higher silicon content. It is also seen from fig. 8 that α/β decreases with decreasing

frequency; a lower β does in fact therefore increase the second as well as the first term on the right-hand side of (9).

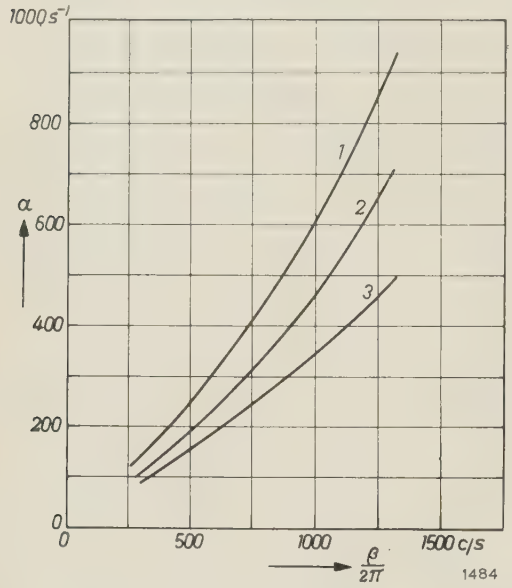


Fig. 8. Damping factor α as a function of frequency $\beta/2\pi$ of the transient, for various types of silicon steel used for the core laminations of a 2 × 4 kV, 25 mA transformer. Curve *1*: low Si content, laminations 0.5 mm thick. Curves *2* and *3*: higher Si content, laminations respectively 0.5 mm and 0.35 mm thick.

For the above type of transformer, loaded with several neon tubes in series (13 mm diameter, total length 6 m), *fig. 9* shows the measured "critical mains voltage" in relation to the nominal mains voltage as a function of the total stray capacitance *C*. (By critical mains voltage is meant the lowest voltage at which the tubes burn steadily, that is without flickering; it is thus the mains voltage at which the generated reignition voltage is only just sufficient.) The dashed curve was calculated from

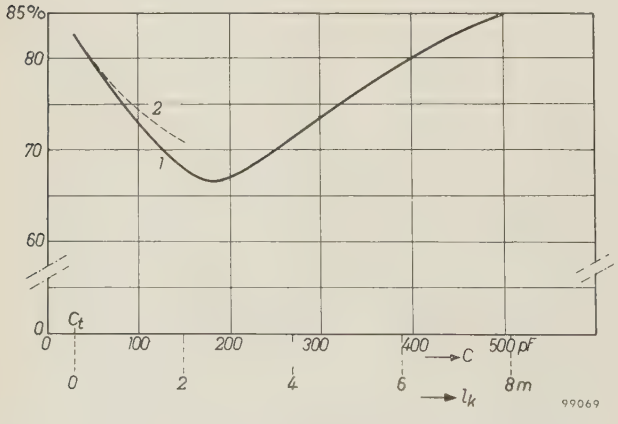


Fig. 9. Ratio of critical to nominal mains voltage, as a function of the stray capacitance *C*. Curve *1*: measured. Curve *2*: calculated. Of the stray capacitance 30 pF was contributed by the transformer; the rest was due to the cable, whose lengths l_k correspond to the abscissa values. Transformer 2×4 kV, 25 mA.

equation (9). Its deviation from the measured curve is mainly attributable to the non-linearity of the inductance *L*. The fact that the measured curve rises again on the right is due to the strong current surges affecting the steady burning of the tube; we shall return to this subject presently.

Since the mains voltage may be lower than the nominal value, and since the ignition voltage may increase during the life of neon tubes, the requirement for *new* tubes in the most *unfavourable* arrangement (with very short connections) is that the critical mains voltage should not exceed 80% of the nominal value.

Current surges

At the moment of reignition the stray capacitance *C*, which is in parallel with the neon tubes, has the voltage $u(\tau)$ across it; see equation (9). The reignition makes the tubes abruptly conducting, causing the sudden discharge of the capacitance. The resultant current pulse (*fig. 10*) is of the form:

$$i = I\,e^{-\varepsilon t}.$$

The time constant ε depends on the magnitude of *C* and on the neon tubes connected.

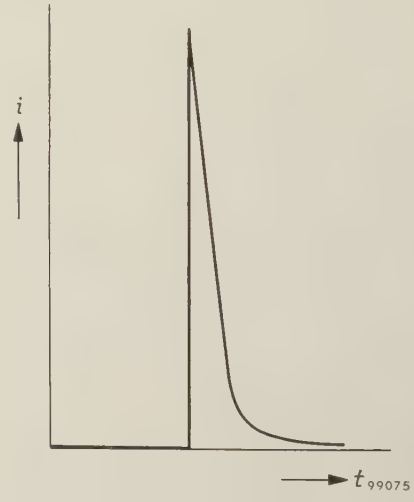


Fig. 10. Waveform of the discharge current *i* of the stray capacitance through the neon tube, on reignition.

The first discharge surge is followed by a series of current surges in increasingly rapid succession; their amplitude decreases whilst the instantaneous value i_s of the steady-state current increases. The explanation of this phenomenon is as follows. The first current surge, starting at $t = 0$ (*fig. 11*), leaves behind in the gas a strong concentration of free charge carriers, which can only disappear by recombination. This concentration, and hence the conductivity of the gas, is for a time so great

that the potential gradient of the discharge is virtually zero. The neon tubes then constitute (during an interval t_1-t_2) a short-circuit across the transformer. During this interval, the current through the tubes is for practical purposes equal to the short-circuit current i_k of the transformer secondary (the time constant ε of the current peaks is small compared with the time constant of the short-circuited secondary).

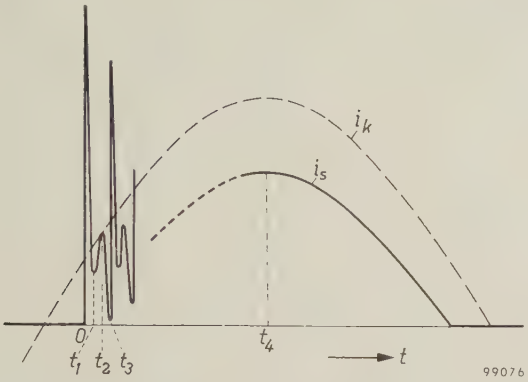


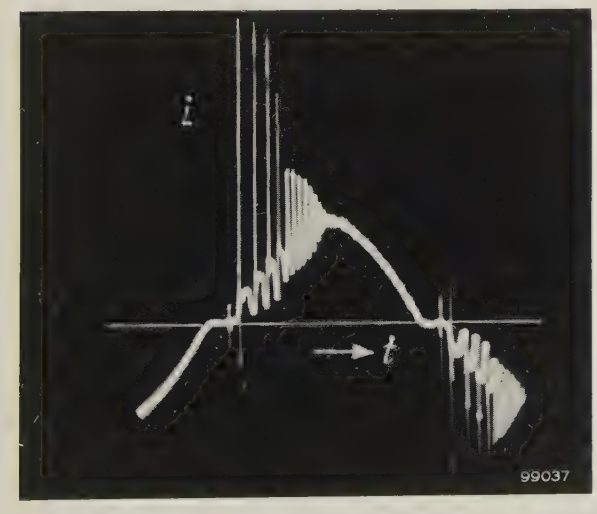
Fig. 11. Waveform of steady-state current i_s through a neon tube, showing superposition of discharge surges. i_k short-circuit current of transformer.

Thereafter the charge-carrier concentration is reduced by recombination. The potential gradient in the gas discharge therefore increases fairly rapidly, so that at the same time the current decreases somewhat (at t_2). The increase in the potential gradient is associated with the charging of the stray parallel capacitance C, but this can only take place at a rate determined by the damping factor α. The effect of this is that the voltage across C cannot increase fast enough, so that the discharge current must decrease still further (beyond t_2) until the discharge extinguishes (at t_3). Not until the potential across C is high enough does reignition take place, whereupon the process is repeated; see the oscillograms in *fig. 12*.

With each repetition the conditions gradually change, because up to the moment t_4 (fig. 11) i_s is still increasing. As a result, the variations in the potential gradient will gradually become smaller and the reignition voltage lower. Moreover, the current required to rapidly charge the stray capacitance becomes relatively less significant. This explains why the current surges become weaker as i_s increases, and follow each other more rapidly; this is to be seen in fig. 12. Provided the stray capacitance is not unduly high, the gas discharge will gradually approach the steady state. However, if C is so high that this state has not yet been reached at $t = t_4$, then, with i_s now decreasing, the current surges will gradually become stronger again and this can cause premature extinction of the gas discharge. This phenomenon is irregular in occurrence, causing the neon tube to burn unsteadily. The voltage oscillogram in fig. 12 shows two waveforms superimposed on one another, which is a consequence of the irregular extinction. The voltage pulse occurring upon the premature extinction is not always of sufficient amplitude to cause reignition. In that case the reignition of the tubes is also irregular. This phenomenon accounts for the upward trend of the right-hand part of the measured curve in fig. 9.

Effect of current surges on the life of neon tubes

During the first current surge an energy of $\frac{1}{2}CU_i^2$ is dissipated, U_i being the reignition voltage. In a



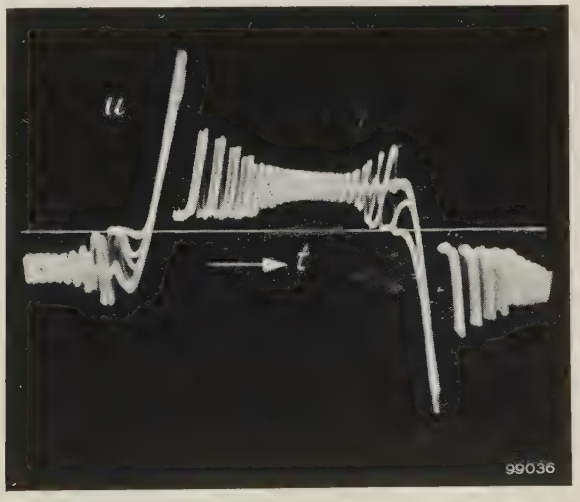


Fig. 12. Oscillograms of the current i through a neon tube and the voltage u across the tube. Owing to irregular extinction, two slightly differing voltage waveforms are seen superimposed in the photograph of u.

practical case (transformer 2×4 kV, 25 mA, which feeds a 6 m long neon tube through a balanced cable of 2×8 m), $C = 500$ pF and $U_i = 12.5$ kV. The energy of the surge is then 0.04 joule. The time in which this energy is dissipated is of the order of 1 microsecond, so that the power of the surge is of the order of 40 kW. Surges of this magnitude have a number of harmful consequences:

1) The tube electrodes are exposed to severe sputtering, which is the main cause of neon being absorbed on the glass wall.
2) The electrodes enter into mechanical vibration (which manifests itself in a high-pitched hum); in the long run this is likely to cause leakage at the glass seals.
3) The surges are a source of radio interference.

We shall deal here only with the effect mentioned under (1); the fact that unduly heavy current surges shorten the life of a neon tube is the primary reason for taking measures to avoid them.

The sputtering process takes place as follows. Particles are dislodged from the electrodes and strike the surrounding glass wall. They entrain a number of gas atoms, and these are absorbed into the layer of metal that forms on the glass. The amount of neon that disappears in this way with each surge may be assumed to be proportional to the average number of metal particles dislodged per surge and proportional to the number of neon atoms N per unit volume. The dislodging of the particles being a process involving an activation energy, the average number of particles dislodged per surge can be put as proportional to the energy P of the current surges per unit time, giving:

$$-V\,\mathrm{d}N = a\,N\,P\,\mathrm{d}t,$$

where a is a proportionality factor. If N_0 is the original number of neon atoms per unit volume, we obtain:

$$\ln\frac{N_0}{N} = a\,\frac{P}{V}\,t.$$

The neon tube has reached the end of its useful life ($t = T$) when N has dropped to a value N_T:

$$\ln\frac{N_0}{N_T} = a\,\frac{P}{V}\,T,$$

whence:

$$T = \left(\frac{1}{a}\ln\frac{N_0}{N_T}\right)\frac{V}{P} = A\,\frac{V}{P}, \qquad (10)$$

where the constant $(1/a)\ln(N_0/N_T)$ is written as A.

On the assumptions made above, the life should be inversely proportional to the energy P of the surges per second, and, if P is constant, directly proportional to the gas volume V; thus, for a given tube diameter and initial pressure, the life is directly proportional to the tube length l_b.

This explains the fact, known from experience, that if two neon tubes of different length are connected in series, the long tube will have a longer life than the short one (the power P is dissipated equally over the four electrodes, at least at the beginning). Series-connected tubes should therefore differ as little as possible in length. This restriction apart, no useful purpose is served by dividing the total tube length into sections. This does not appreciably reduce the reignition voltage, nor therefore the total energy of the current surges. Since a certain loss of energy occurs in the gas discharge at each electrode, the total loss is increased by the greater number of electrodes, and the total tube length must therefore be smaller than in the case of one long tube. A smaller length per tube means a smaller volume of gas, and therefore, according to (10), a shorter life.

The conclusion just drawn, viz. that $T \propto l_b$, is valid for $P =$ constant, which is the case in the comparison of series-connected neon tubes. We now compare two installations consisting of the same number of neon tubes connected in series, but where the tubes in the one installation are n times longer than in the other. In each of the installations the transformer voltage must be matched to the overall tube length, in which case (where the capacitance C is constant) the life T is *inversely* proportional to l_b: since the required reignition voltage is approximately proportional to l_b, and P is thus proportional to l_b^2, it follows from (10) that $T \propto l_b^{-1}$.

If the high-tension cable is fairly long, the transformer capacitance is negligible compared with the cable capacitance C_k, which is proportional to the length l_k of the cable. The power P is then approximately proportional to l_k, so that, where d is the diameter of the neon tubes, it follows from (10) that

$$T \propto \frac{d^2}{l_b\,l_k}. \qquad (11)$$

In the Philips factory at Roosendaal, The Netherlands, life tests have been carried out on neon tubes of different lengths and diameters, and connected by cables of varying lengths. Some results are given in *fig. 13*. Fig. 13*a* shows the average life T as a function of cable length l_k, for tubes of various lengths l_b. Fig. 13*b* shows T as a function of the neon volume V, for various cable lengths l_k. The neon

tubes were connected in series (P therefore being constant). The results are in broad agreement with the formulae (10) and (11). The fact that fig. 13*a* shows the life to be roughly inversely proportional to l_k, and therefore inversely proportional to the cable capacitance, supports the premise assumed earlier that the average number of metal particles dislodged per surge is proportional to the energy P of the surges per unit time.

From (10) and (11) one would expect to find in fig. 13*a* and *b*, in which both scales are logarithmic, straight lines having a slope of exactly $-45°$ and $+45°$, respectively. The fact that the slopes actually found differ somewhat from 45° is probably attributable to the fact that, owing to the series-connection of tubes of different lengths, the gas disappeared most rapidly from the shortest tube; the shorter the tube, the more rapidly the gas pressure declined. After some time the pressures will have become different and the energy of the current surges will no longer have been uniformly distributed over the electrodes.

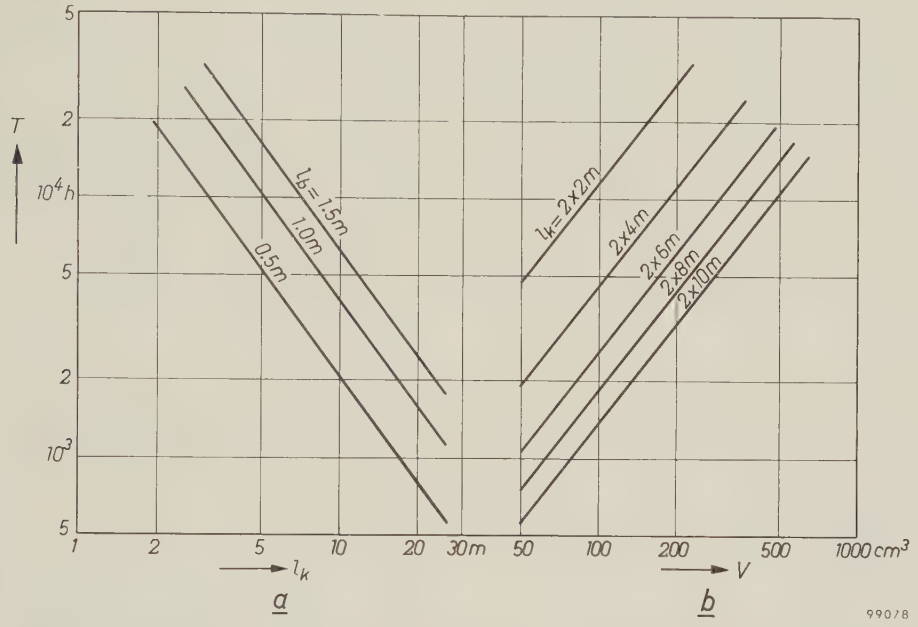


Fig. 13. Results of life tests on neon tubes carried out in the Philips factory at Roosendaal. *a*) Average life T as a function of cable length l_k, for tubes of various lengths l_b. *b*) T as a function of gas volume V for series-connected tubes, for various cable lengths l_k. All tubes used for these tests had the same cross-section.

According to (11) it is possible with shorter tubes (and appropriately smaller transformers) to obtain a longer life or, for the same life, one can use a longer cable. This is illustrated by some examples in *fig. 14a-d*. Assuming that the distance AB from the transformer to the one end of the neon tubes is the same in these four examples, it can be seen with the aid of the above considerations that cases *a* and *b* are completely equivalent, both as regards life and cable length. In case *c* we may expect the life to be about 1.5 times longer, the cable length required being also 1.5× longer. In case *d* the life is about twice as long, and the cable length also twice as long. The voltage on the cable is much lower for *c* and *d* than for *a* and *b*.

Surge limitation

The foregoing considerations have shown clearly the desirability of finding some means of limiting the current surges and of dissipating the energy of the charged stray capacitances as far as possible *outside* the neon tubes.

The surges can be limited by means of a resistor or a choke. Being cheaper, resistors are frequently used for this purpose. The method is to connect the resistor to the end of the cable in series with the neon tubes. The effect is slight, however, since the value of the resistance R_1 must be restricted to about 10% of the load resistance represented by the neon tubes; otherwise the loss in the resistance would demand too great a sacrifice in the length of tube in relation to the transformer voltage. The current peaks are then reduced by a factor of about 2, which is hardly adequate.

If the current surges are limited by a choke (inductance L_1, resistance R_1) at the end of the cable, the current surge is given approximately by:

$$i = \frac{E}{L_1} \exp\left(-\frac{R_1}{2L_1}\, t\right) \frac{\sin qt}{q},$$

where q is given by

$$q = \sqrt{\frac{1}{L_1 C} - \frac{R_1^2}{4L_1^2}}.$$

q may be either real or imaginary. The special case, $q = 0$, leads to:

$$i = \frac{E}{L_1} t \exp\left(-\frac{R_1}{2L_1} t\right).$$

The peak value of i is then $2E/eR_1 = 0.74\ E/R_1$, and thus little less than in the case where only a resistance R_1 is used. Imaginary values of q are therefore useless here. In order to examine the possibilities presented by the use of real values of q (i.e. periodic solutions), we consider the case

$$\frac{1}{L_1 C} \gg \frac{R_1^2}{4L_1^2}.$$

The peak value I of the current i is then

$$I \approx E \sqrt{\frac{C}{L_1}}.$$

It is evident that in this way a fairly small inductance L_1 (provided its own capacitance is small enough) is able to reduce the current surges by a factor of, say, 10. The condition is, however, that the oscillations produced in the L_1-C circuit should be sufficiently damped. This is necessary, 1) in order to dissipate most of the surge energy outside the neon tubes, and 2) because otherwise the tubes would again burn unsteadily, for the same reason as described above.

The oscillations can be damped either by shunting a resistance across the coil, or by ensuring that the iron losses in the core of the coil are sufficiently high.

Although the unfavourable effects of high cable capacitances can be appreciably reduced in this way, when long cables are necessary it is advisable in addition to use low-capacitance cables. For example, the conventional lead-sheathed cable, which has a capacitance of 120 pF per metre, can advantageously be replaced by a rubber-insulated wire enclosed in 5/8″ piping, which has a capacitance of only 65 to 70 pF per metre.

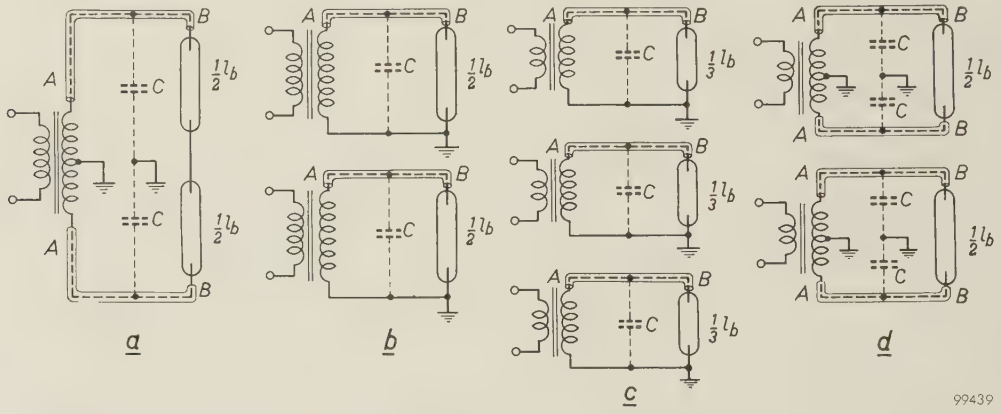


Fig. 14. Various circuits for the supply of neon tubes. The total length l_b of the tubes is the same in all four cases. Similarly, the distance AB is assumed to be the same in all four cases.
a) Two tubes, each $\frac{1}{2}l_b$ in length, fed by a balanced transformer.
b) Each tube fed by its own transformer. Equivalent to (*a*).
c) Three tubes, each $\frac{1}{3}l_b$ in length, each fed by its own transformer. The life is about 1.5× that in (*a*) and (*b*). The required length of cable is also about 1.5× that of (*a*) and (*b*).
d) Two tubes, each $\frac{1}{2}l_b$ in length, each fed by its own balanced transformer. Here the life is about 2× that of (*a*) and (*b*), and the required length of cable is also twice as long.

Unbalanced cables

Where unbalanced transformers are used (one side of the secondary being earthed, see figs. 2*a* and *c* and 14*b* and *c*), the effect of cable capacitance can be limited by connecting the unearthed secondary terminal by as short a lead as possible to the neon tube; the other (earthed) lead may be as long as required.

In the case of balanced transformers (where the centre of the secondary winding is earthed, see fig. 14*a* and *d*) it is possible to limit the current surges by unbalanced connection of the cables, that is by making one high-tension cable much shorter than the other. During the surge the two stray capacitances are effectively in series for the discharge; if the one capacitance is much smaller than the other, the equivalent capacitance is virtually equal to the smaller capacitance, which therefore mainly determines the strength of the surges. One might thus be inclined to make the short cable extremely

short, but in that case, if the other connection is long (> 10 m), complex transients may arise after the reignition, and these will almost immediately extinguish the discharge. However, with a length ratio of 1 : 10 this effect is not troublesome, and moreover the life of the neon tubes is appreciably lengthened in this way without other measures being necessary.

Finally, some remarks on the maximum voltage appearing across the cables. When the installation is operating, this voltage in the case of unbalanced transformers is equal to the reignition voltage, and in the case of balanced transformers it is equal to half the reignition voltage. In practice this amounts in both cases to a peak voltage of at the most 6.5 kV. When the transformer is not loaded (neon tubes not burning) the maximum voltage on the cable is $1/(1-\omega^2 LC)$ times higher than with an infinitely short cable. Where long cables are used, this means an increase of 10 to 15% above the no-load voltage that would be present with an infinitely short cable.

Summary. The high voltage required for the reignition of neon tubes would call for relatively expensive power transformers if use were not made for this purpose of the transient produced periodically when the current has dropped to zero. This transient consists of a damped oscillation of a frequency (500-1200 c/s) partly determined by stray capacitances. The height of the first positive voltage peak governs the length of tube that can be connected to a given transformer; it is affected by the quality of the laminations of the transformer core, by the magnitude of the magnetic induction in the core, and to a certain extent by the capacitance (hence the length) of the high-tension cable(s) between transformer and neon tube.

With each periodic reignition of the neon tubes the stray capacitances discharge through the neon tube. These strong current surges cause the tubes to burn unsteadily and shorten their useful life. The surges can be limited by constructing the installation from a number of smaller units, by using unbalanced high-tension cables of low capacitance, and by connecting a small inductance in series with the neon tubes.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS BY THE STAFF OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to Philips Electrical Ltd., Century House, Shaftesbury Avenue, London W.C. 2.

2681: J. Halberstadt: Some experiments with radioactive preparations of 2,4,5,4′-tetrachlorodiphenylsulphone, a new acaricide (Meded. Landbouwhogesch. Opzoekingsstat. Gent **23**, 788-794, 1958, No. 3/4).

Radioactive-isotope tracer technique has been used to study the metabolism of "Tedion V18" (an acaricide against mites) in plants and animals. The active part of Tedion, viz. 2,4,5,4′-tetrachlorodiphenylsulphone, was labelled by partly replacing the sulphur with the isotope ^{35}S. Apple trees were sprayed with the labelled preparation, formulated as a wettable powder and as a miscible oil. From measurements of the radioactivity, it was found that an active residue remains on the leaves for a long time; the Tedion that permeates in the leaf is depleted by chemical change and transport but is replaced by constant permeation from outside. Miscible oil appears to be a more suitable formulation than wettable powder. Administered to rats (a maximum of 100 mg per kg body weight), 40-45% of a Tedion dose is found unchanged after 48 hours in the faeces; the remainder is broken down.

The substance of this paper will shortly appear in the Philips Technical Review.

2682: W. Duyfjes: Some problems in pesticide formulation (Meded. Landbouwhogesch. Opzoekingsstat. Gent **23**, 837-845, 1958, No. 3/4).

Discussion on the formulation of pesticides, i.e. chemicals used against insects, mites, fungi and weeds. Particular attention is paid here to pesticides prepared for spraying with water. It is shown that for each pesticide a separate investigation is necessary to determine the best formulation. Not only must the physical and chemical properties be considered (e.g. the stability of suspensions or emulsions), but the formulation must be tested for its biological effectiveness. Sometimes formulations with the best physico-chemical characteristics give less favourable biological results, and vice versa. Examples are given by way of illustration. See also Philips tech. Rev. **19**, 165-176, 1957/58.

2683: J. H. Stuy: Nucleic acid synthesis in ultraviolet-irradiated Bacillus cereus (J. Bacteriol. **76**, 668-669, 1958, No. 6).

Further investigations into the function of nucleic acids in physiological processes in bacteria (see also No. **2635** of these Abstracts). Kelner has shown that U.V. irradiation inhibits the

formation of deoxyribonucleic acid (DNA) by the organism. Some time after stopping the irradiation, the formation of DNA can once more begin. This paper concerns an investigation into the composition of DNA and RNA (ribonucleic acid) as these are formed after irradiation in *Bacillus cereus* p2. With the somewhat crude chemical methods used, no difference could be detected between the DNA and the RNA formed before the irradiation. To what extent the post-irradiation DNA differs biologically from that normally formed, is a problem that will be further investigated.

A 6: E. Baronetzky: An apparatus, capable of being heated, for high-vacuum grinding of solids (J. sci. Instr. **35**, 427-428, 1958, No.11).

A rotary file is mounted on one end of a spindle free to rotate in glass-metal bearings. At the other end of the spindle is a steel rotor with a layer of copper brazed to its periphery. This whole system together with the material to be ground and the feed guides are surrounded by an envelope of hard glass and can be heated to 400 °C. The drive is via a multi-poled magnetic ring of ferroxdure, concentric with the rotor and outside the glass envelope, itself driven by a variable speed DC motor.

A 7: A. Rabenau and P. Eckerlin: $BeSiN_2$, eine neue Verbindung mit Wurtzitstruktur (Naturwiss. **46**, 106-107, 1959, No. 3). ($BeSiN_2$, a new compound having a wurtzite structure; in German.)

Note concerning a new compound $BeSiN_2$, one of the group of $A^{II}B^{IV}X^{V}_2$. It has the wurtzite structure analogous to the nitrides of Al, Ga and In, and is colourless. The crystallographic data are: $a = 2.872 \pm 0.004$ Å, $c = 4.674 \pm 0.004$ Å. The X-ray density was found to be 3.24 g/cm³; the value from weighing and volume determination was 3.12 g/cm³.

A 8: H. G. Grimmeiss and H. Koelmans: Über die Kantenemission und andere Emissionen des GaN (Z. Naturf. **14a**, 264-271, 1959, No. 3). (On edge emission and other kinds of emission of GaN; in German.)

The fluorescent properties of pure GaN and GaN doped with Li, Zn and Mg have been measured under U.V. radiation and cathode rays as functions of the temperature. Several emission bands have been found. The shortest-wavelength band is attributed to recombination of free electrons with free holes.

A 9: F. Kettel: Die Wärmeleitfähigkeit von Germanium bei hohen Temperaturen (Phys. Chem. Solids **10**, 52-58, 1959, No. 1). (The thermal conductivity of germanium at high temperatures; in German.)

An account is given of thermal conductivity measurements on germanium single crystals, both intrinsic and with high hole conductivity. The author took mobility values, and values for the width of the forbidden zone and its dependence on temperature — all necessary for theoretical discussion of the electric component of thermal conductivity — from the measurements of other authors, and confirmed them by comparison with his own measurements of thermoelectric force and electrical conductivity. The measured values of the electronic component of thermal conductivity, of both intrinsic and hole-conductive germanium, are about 4 times as high as those determined theoretically.

A 10: A. Goertz: Zur Theorie der diffusen Reflexion und Transmission beim Vorliegen elastischer Vielfachstreuung (Z. Physik **155**, 263-274, 1959, No. 3). (On the theory of diffuse reflectivity and transmittivity in case of elastic multiple scattering; in German.)

In order to determine the diffuse reflectivity and transmissivity of multiply scattered particles, there is not necessarily the need of solving the complete differentio-integral equation of multiple scattering. It can be shown that the problem reduces to solving a set of two simultaneous differential equations of the first order, provided that the amount of sideward scattering may be neglected. The solutions of these equations furnish all data required, if the mechanism of single scattering is known.

A 11: E. Kauer, O. E. Klinger and A. Rabenau: Über Leitfähigkeitsmessungen im System ZrO_2-MgO (Z. Elektrochemie **63**, 927-936, 1959, No. 8). (Electrical-conductivity measurements on the system ZrO_2-MgO; in German.)

The electrical conductivity of pure ZrO_2 and of the system ZrO_2-MgO has been measured as a function of the temperature between 500 and 1500 °C. Further, thermoelectric measurements have been carried out, which indicate positive charge carriers. The measurements indicate that the conductivity (at least at the partial pressure of oxygen in the air) is preferentially due to oxygen vacancies. It is shown that, with a knowledge of the conduction mechanism, conductivity measurements as a function of time and temperature give a sensitive method of following the solid-state reaction involved. It not only gives a picture of the phase concentrations but also gives some idea of the progress of the transformation with time.

R 376: K. Jost: Impedance transformations through lossless two-ports represented by fractional linear transformations of the unit circle (Philips Res. Repts. **14**, 301-326, 1959, No. 4).

Any impedance transformation by a lossless two-port network corresponds to an automorphism of the unit circle in the complex reflection-coefficient plane. Such a mapping of the unit circle, determined by three parameters, can always be resolved into three successive elementary transformations: a rotation about the origin, a hyperbolic transformation corresponding to the impedance transformation by an ideal transformer, and a further rotation about the origin. Resolution in this way leads to a simple graphical method for the determination of the transformed impedance and for the treatment of cascades of lossless two-ports. Cascades giving an ideal transformer are investigated. The determination of the parameters of a two-port (calibration of the two-port) is described.

R 377: M. T. Vlaardingerbroek: Noise in electron beams and in four-terminal networks (Philips Res. Repts. **14**, 327-336, 1959, No. 4).

It is shown that the calculation of the minimum noise figure of electron-beam amplifiers, obtainable by varying the physical properties of the lossless beam-transducer between the electron gun and the interaction region, can be performed in a manner which is similar to the calculation of the minimum noise figure of network fourpoles based on the method of varying the signal-source impedance. Further, the similarity between the propagation of noise fluctuations along electron beams and the transformation of noise sources across passive fourpoles is emphasized.

R 378: P. Penning: Rate of diffusion-limited annihilation of excess vacancies (Philips Res. Repts. **14**, 337-345, 1959, No. 4).

The rate of removal and the spatial distribution of excess vacancies are calculated for the case where the transport to vacancy sinks takes place by diffusion and where dislocations in the volume of the sample, the surface of the sample, or both may act as sinks. The medium in which the diffusion takes place is considered as a continuum. It is found that in the case where the dislocations are the only sinks, the decay in average concentration is exponential. The decay time constant is almost inversely proportional to the dislocation density. In the case where the surface is the only sink an appreciable deviation from an exponential decrease in average concentration does occur in the beginning of the removal process. In a bar of square cross-section the contribution of the two types of sink is about equal if a few tens of dislocations emerge through the smallest cross-section.

R 379: H. U. Harten: Surface recombination of silicon (Philips Res. Repts. **14**, 346-360, 1959, No. 4).

The surface-recombination velocity of electrons and holes in silicon is investigated by measuring the photovoltaic effect of a p-n-junction alloyed on a thin silicon wafer. The method has been reported previously; its principle is explained here on the basis of a hypothetical experiment. It appears from the measurements that after treating the surface with an aqueous solution of potassium dichromate the recombination is lowered by ozone in the ambient atmosphere and raised by moisture. The opposite behaviour is observed after etching with hydrofluoric acid. In principle there is no difference in the behaviour of n-type and p-type silicon. In many cases the surface recombination can be decreased with light, particularly if the surface treatment tends to form an inversion layer. The observations can be interpreted by the assumption that the surface recombination on silicon is due to recombination centres and therefore influenced by the position of the Fermi level at the surface, and that an additional influence is due to the voltage across the surface barrier layer.

R 380: D. de Nobel: Phase equilibria and semiconducting properties of cadmium telluride (Philips Res. Repts. **14**, 361-399, 1959, No. 4; continued in No. **R 384**).

In this thesis (Leyden, May 1958) the relation is studied between the electrical and optical properties of single crystals of cadmium telluride and the conditions of preparation. The p-T-x diagram of the system cadmium-tellurium is described, showing the temperatures and Cd pressures at the maximum melting point and at the melting point of stoichiometric CdTe. The compound is purified by zone refining; foreign atoms are incorporated by zone levelling and single crystals are grown, which are reheated at various Cd pressures between 700 and 1000 °C and then quenched to room temperature. On these samples conductivity and Hall measurements are performed at various temperatures, which lead to values of the concentration of charge carriers, of the ionization energies of the various centres and — for n-type samples with shallow donors — of the density-of-state effective mass of the free electrons ($m_n^*/m = 0.14 \pm 0.04$). Thermoelectromotive-force measurements of various n and p-type samples

lead to values of the effective mass for both types of carrier, depending on the value adopted for the transport energy of electrons and holes (m_n^*/m = 0.13-0.066; m_p^*/m = 0.41-0.22). From a hydrogen-like model for the shallow donor, an inertial effective mass for the electrons of $(m_n^*)_i = 0.147$ is obtained. From optical transmission measurements at various temperatures the band gap is found to be 1.50 eV at room temperature and the temperature dependence 2.34×10^{-4} to 5.44×10^{-4} eV/°K. Peaks in the spectra of photoluminescence and photoconductivity can be correlated with a band-band transition and with transitions between levels caused by known centres, and one of the bands. A theoretical discussion is presented of the various equilibria which determine the state of CdTe at high temperatures. By assuming a certain band scheme and certain values for the equilibrium constants, it is possible to calculate the concentrations of charge carriers and centres at room temperature as a function of the Cd pressure at which the crystals were reheated. A qualitative comparison between the experimental and various theoretical diagrams leads to the adoption of a definite band scheme, in which two levels are attributed to the cadmium vacancy and one each to interstitial cadmium, to indium (as a specific donor) and to gold (as a specific acceptor). From a quantitative discussion of the diagrams the values of some equilibrium constants are obtained, viz. the Frenkel constant K_F and the reduction constant K_r, describing the equilibrium: crystal-vapour. The temperature dependence of these constants leads to the activation energies required for the atomic processes involved. Finally, the association effects which take place during the quenching of activated samples are analysed. In the appendix a discussion is given of the type of bonding in CdTe.

R 381: Joshua Ladell and William Parrish: Determination of spectral contamination of X-ray tubes (Philips Res. Repts. **14**, 401-420, 1959, No. 5).

An X-ray method is outlined for the qualitative and quantitative analysis of the spectral purity of X-ray tubes. The method, based on well-known principles, employs a standard diffractometer equipped with a xenon-filled proportional counter and molybdenum-foil analyser. The various theoretical and practical aspects of the determination of the correction factors for the comparison of different wavelengths are described.

R 382: S. Duinker: Generalization to non-linear networks of a theorem due to Heaviside (Philips Res. Repts. **14**, 421-426, 1959, No. 5).

A theorem enunciated by Heaviside and proved by Lorentz for linear electromagnetic systems subjected to a constant electric force which is suddenly impressed, is extended to electrical networks comprising non-linear reactances and linear dissipances. The theorem states that the amount of work to be done by a constant-voltage source exclusively to sustain the transients arising as the network changes from the rest state to the steady state, is equal to the excess of the sum of electric energy and co-energy over the sum of magnetic energy and co-energy.

R 383: J. F. Marchand and A. Venema: Note on superconducting tantalum films (Philips Res. Repts. **14**, 427-429, 1959, No. 5).

Thin superconductive tantalum films are difficult to produce. In this note it is shown how such films can be made by evaporation under extremely low pressure. The films thus prepared have been found to exhibit practically the same superconducting properties as the bulk material.

R 384: D. de Nobel: Phase equilibria and semiconducting properties of cadmium telluride (Philips Res. Repts. **14**, 430-492, 1959, No. 5). Continuation of **R 380**.

Now available:

R. van der Veen and G. Meijer: Light and plant growth (Philips Technical Library 1959, pp. 164, 92 illustrations).

The first five chapters of this book are devoted to the various processes in plants that are caused or influenced by light; the sixth chapter deals with the practical application of this knowledge, in particular with regard to the irradiation of plants by artificial light. The titles of the chapters are: I. Light measurement for plant irradiation, II. Photosynthesis, III. Phototropism, phototaxis and photonasty, IV. Photoperiodism, V. Effect of the colour of the light, VI. The use of artificial light in horticulture.